

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 23, 2004, 11:26:23 ; Search time 7464 Seconds
(without alignments)
10994.736 Million cell updates/sec

Title: us-09-701-572-1
Perfect score: 2006
Sequence: 1 gattcgacaggaagagaaa.....agaaaaaaaaaaaaaa 2006

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.un.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2006	100.0	2006	6	AX008931	Sequence
2	1986.4	99.0	1988	8	AF079404	Medicago
3	719.8	35.9	1541	8	AY128834	Arabidops
4	719.8	35.9	1865	8	AY095885	Arabidops
5	708.4	35.3	1459	8	AY091235	Arabidops
6	708.4	35.3	1511	8	AY063875	Arabidops
7	581.8	29.0	8450	8	AF134835	Medicago
8	472.4	23.5	1511	8	BT002165	Arabidops
9	472.4	23.5	1891	8	AY095881	Arabidops
10	405.2	20.2	3033	5	BC045037	Xenopus lae
11	403.6	20.1	2649	5	XLNAPFRP	BC048038 Danio rer
12	357	17.8	2089	5	BC048038	BC006616 Mus muscu
13	352.8	17.6	1861	10	BC006616	AF083809 Mus muscu
14	351.2	17.5	2258	10	AF083809	AB013452 Homo sapi
15	340.8	17.0	1491	9	AB013462	BT007115 Homo sapi
16	339.2	16.9	1482	9	BT007115	BT008152 Synthetic
17	339.2	16.9	1482	12	BT008152	BC013413 Homo sapi
18	339.2	16.9	1866	9	BC013413	AF083810 Homo sapi
19	339.2	16.9	2291	9	AF083810	AB033068 Homo sapi
20	339.2	16.9	5008	9	AB033068	BT001403 Drosophill
21	334.8	16.7	3232	3	BT001403	AF421546 Gallus ga
22	331.8	16.5	1362	5	AF421546	AF080397 Homo sapi
23	331.2	16.5	1666	9	AF080397	Y14162 Drosophilla
24	330	16.5	2605	3	DMENAPRP	AF421548 Gallus ga
25	317.8	15.8	1488	5	AF421548	AX489295 Sequence
26	303.8	15.1	1794	6	AX489295	AB005589 Schizosac
27	291.6	14.5	2627	8	AB005589	AL132675 S.pombe c
28	291.6	14.5	41101	8	SPAC144	AB013463 Homo sapi
29	282.4	14.1	1215	9	AB013463	AF433157 Homo sapi
30	280.8	14.0	1215	9	AF433157	AY118173 Ustilago
31	277.2	13.8	2240	8	AY118173	AL355941 Schizosac
32	277	13.8	26756	8	SPEC1198	U33010 Schizosacch
33	277	13.8	37596	8	SPU33008	AL034563 S.pombe c
34	277	13.8	43325	8	SPBC660	U33010 Schizosacch
35	275	13.8	85837	8	SPU33010	AX654709 Sequence
36	259	12.9	46607	2	ACU13865	AC013865 Drosophill
37	259	12.9	160629	3	AC023715	AC023715 Drosophill
38	259	12.9	310990	3	AE003431	AE003431 Drosophill
39	259	12.4	1973	3	AF052575	AF052575 Spisula s
40	248.2	12.4	1441	3	AY051537	AY051537 Drosophill
41	248	12.4	1441	3	AY051537	AC020529 Drosophill
42	248	12.4	44588	2	AC020529	AC004642 Drosophill
43	248	12.4	14832	3	AC004642	AC099018 Drosophill
44	248	12.4	154840	3	AC099018	AE003462 Drosophill
45	248	12.4	303823	3	AE003462	

ALIGNMENTS

RESULT 1	AX008931	Sequence	1 from Patent WO9964451.	2006 bp	DNA	linear	PAT 06-SEP-2000
LOCUS	AX008931	Sequence	1 from Patent WO9964451.				
DEFINITION	AX008931	Sequence	1 from Patent WO9964451.				
ACCESSION	AX008931	Sequence	1 from Patent WO9964451.				
VERSION	AX008931.1	GI:9996328					
KEYWORDS		Medicago sativa					
SOURCE		Medicago sativa					
ORGANISM		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;					
		rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;					
		Medicago.					
REFERENCE	1						
AUTHORS		Cebolla,A., Kondorosi,E. and Kondorosi,A.					

QY	1021	ACAGAAGATTTGTTAGTAACTGTGAGGACACAAATCAGAGGTTTGTGACTGAAGTG	1080
Db	1021	ACAGAAGATTTGTTAGTAACTGTGAGGACACAAATCAGAGGTTTGTGACTGAAGTG	1080
QY	1081	GTCTATCATTAACCGTGAAGTTGGGATCTGGAGGAAATGACACAAATTTGTTTGGAA	1140
Db	1081	GTCTATCATTAACCGTGAAGTTGGGATCTGGAGGAAATGACACAAATTTGTTTGGAA	1140
QY	1141	TCAACATCAACCCAGCGTCTCTCAAGTACTGTGAGGACACACAGAGCTGTTAAAGCTAT	1200
Db	1141	TCAACATCAACCCAGCGTCTCTCAAGTACTGTGAGGACACACAGAGCTGTTAAAGCTAT	1200
QY	1201	TGCTAGGTTCTCTCATCTTCATGAGTCTTCTGCACTCTGGAGGAACTGAGATAGATG	1260
Db	1201	TGCTAGGTTCTCTCATCTTCATGAGTCTTCTGCACTCTGGAGGAACTGAGATAGATG	1260
QY	1261	TATTCGTTTTTGGAAATACACCAAACTCACACCTTAGCTGTATGGACATGGAAGTCA	1320
Db	1261	TATTCGTTTTTGGAAATACACCAAACTCACACCTTAGCTGTATGGACATGGAAGTCA	1320
QY	1321	GGTTTGCATCTGTCTGCTCCAAAATGTCAACGAATAGTAAGCACACATGGTACTC	1380
Db	1321	GGTTTGCATCTGTCTGCTCCAAAATGTCAACGAATAGTAAGCACACATGGTACTC	1380
QY	1381	CCAGAACCAAGATTTATTTGTTGGAGTACCCCACTATGTCAGAGCTGGGACTTTACCGG	1440
Db	1381	CCAGAACCAAGATTTATTTGTTGGAGTACCCCACTATGTCAGAGCTGGGACTTTACCGG	1440
QY	1441	CCATCTTATAGGTTCTCTATCTTCCCATCTCCAGATGACAGACTATTGTAACCTGG	1500
Db	1441	CCATCTTATAGGTTCTCTATCTTCCCATCTCTCCAGATGACAGACTATTGTAACCTGG	1500
QY	1501	AGCTGAGATGAAACGCTTAGGTTCTGGAATGTTTCCCTTCCCTTAATACAGAAATAC	1560
Db	1501	AGCTGAGATGAAACGCTTAGGTTCTGGAATGTTTCCCTTCCCTTAATACAGAAATAC	1560
QY	1561	TGAAGTGAATCGGAGCAATATCTCTGGAGCACTCTATCAGTGAATGATCTCTGGC	1620
Db	1561	TGAAGTGAATCGGAGCAATATCTCTGGAGCACTCTATCAGTGAATGATCTCTGGC	1620
QY	1621	GTTCAGCCCAATCATGTGGCATATTTCTAAGTTTGGGTTGCTGTGTAGAACTAAATTC	1680
Db	1621	GTTCAGCCCAATCATGTGGCATATTTCTAAGTTTGGGTTGCTGTGTAGAACTAAATTC	1680
QY	1681	TGAGCGGAGAACCACTAGTGGAAAACTTGAATATAAACAACCAACCAAGTAGCATC	1740
Db	1681	TGAGCGGAGAACCACTAGTGGAAAACTTGAATATAAACAACCAACCAAGTAGCATC	1740
QY	1741	TTTACCAACTGGGAGAGCCTTGGAGGAGCTATAAAGTTTTCATATGGCTGCGGTGAT	1800
Db	1741	TTTACCAACTGGGAGAGCCTTGGAGGAGCTATAAAGTTTTCATATGGCTGCGGTGAT	1800
QY	1801	ATTCTCGCATCATGTGTAGTCTCATTTTATATGAAAGATGATACAAATGGGTAAT	1860
Db	1801	ATTCTCGCATCATGTGTAGTCTCATTTTATATGAAAGATGATACAAATGGGTAAT	1860
QY	1861	TATTTCTCTGCACTTATACATGCACTGATGAGTTGTAGCAAGTTTATTAATCTCTT	1920
Db	1861	TATTTCTCTGCACTTATACATGCACTGATGAGTTGTAGCAAGTTTATTAATCTCTT	1920
QY	1921	TTTTTCT	1980
Db	1921	TTTTTCT	1980
QY	1981	TAAACAGAG 1988	
Db	1981	TAAACAGAG 1988	

RESULT 3
AY128834
LOCUS

BASE COUNT 422 a 312 c 360 g 447 t

ORIGIN

Query Match 35.9%; Score 719.8; DB 8; Length 1541;
Best Local Similarity 74.2%; Pred. No. 5.5e-158;

Matches 974; Conservative 0; Mismatches 312; Indels 27; Gaps 4;

Qy 306 CCTCACCTTCGAAAGAACTACTCCGATAGGTTCACTCCGATAGATCTGCTTCCGAAT 365

Db 158 CATCACTATCAAGGTCTATATCTCTGATAGATTTATCCCCAGTAGATCCGATCCAAAT 217

Qy 366 TCGTTTGTGATATCAATTAATCTCCGAC-----AGAAGACGCGATGATAGTTCCAGCG 419

Db 218 TCGTCTTTTCGATCTATCTCTCTCTAGTAAAGATGTAAGGAAGATGGAGCTGGCT 277

Qy 420 CTTATAGCATCTTCTCGAAGACGGCGTGTGTTGACCGGATGTTGCCGTCGGTTACGC 479

Db 278 CTTACGCTACTCTGTTGCGTCCGCGATGTTGCTGCTGAG-----ACGC 322

Qy 480 CGGAAAAACCGACTCGCGCTGATGACATTCGCGAATAGGAATATTTTATAGTATAAGA 539

Db 323 CGGAGAGAGAGATTAATCTAGGTTTCTCTTCC---AGGAATATTTTATAGTTTAA 379

Qy 540 CGGAGAGAGAGATGATCGATGCACTCGCTTTCGCGCTTATGATGATGATTTTGTCTTG 599

Db 380 CGGAGACTCATCGGCTTTTGAATTCGTTTCTCTCTTTCGTTGTTGATGATGATCTCCCTG 439

Qy 600 GTGTTAATCATATA---GTCCGGTTAAGCTCTAGGAAAGTTTCTCCGATCGCTTATAAGG 656

Db 440 GTGTTTCTCATAGTGTCTCTGTTAAGCTCCAGAAAGTCCCGGATCGCGCTATAAGG 499

Qy 657 TTTTGGATGCACTGCTTTGCAAGATGATTTTATCGAATCTGGTATGATGTTCTTCAC 716

Db 500 TATTGGATGCACTCGGCTTTCGAAGATGATTTTATCTGAATCTTGTGGATTTGCTGCAC 559

Qy 717 ACAATGTCTTGGCTGTTTGGTAACTGCTATCTCTCGAATCTTGTAGCAGCA 776

Db 560 AAATGTTCTAGCAGTGGGACTAGGAACTGTGTATTTATGAATGCTTGTAGCAGCA 619

Qy 777 AGGTAACATAATTAATGTAATTTGGGGTTGATGATGTTGTTGTTGTTGGGCTC 836

Db 620 AGGTTACTAAGTATGATGATCTCGAGCTGAGGATAGTGTGTTGCTCACTGGGTTGGCGT 679

Qy 837 AACGTGTTACTCATCTCTGTTGGAATCAATGTTAAAGTTTCAATTTGGGATGAGCAG 896

Db 680 TAGGTGGAACATCATCTGCTGTGTAATAGTACCGGAAAGTTTCAGATAGGATGGGT 739

Qy 897 CAAGATGCAAGAAAGATAAGTCAATGAGGAGCCATCGGTTACGTGTCGGGCTTGGCCCT 956

Db 740 CACGCTGCAAGAGACAGAAATGGAAGTCAATCGTCTAAGAGTTTGGAGCCCTGGCAT 799

Qy 957 GGAGTTCACTCTTTTGTCTTCTGTTGCAAGGATGATGATTTATCAACGAGATATAC 1016

Db 800 GGGGTTTCATCGGTTCTGTCATCTGTAGCAGACAGAGATATCTTTCAGAGAGACATAA 859

Qy 1017 GCACACAGAGATTTTGTAGTAACTGTGAGGACACAAATCAGAGTTTGTGGACTGA 1076

Db 860 GGTGTCAAGAGATCATGTCAGTAAATTTGGCAGGTCTATAATCTGAGGTATGCGACTCA 919

Qy 1077 AGTGGTCAATGATAACCGTGAAGTTGGCATCTGGAGGAAATGACAAATTTGTTGTT 1136

Db 920 AGTGGTCTTATGACAAACAGAGAGTACATCTGTGTGAAACACCAATAGGCTTTTGTAT 979

Qy 1137 GGAATCAACACTCAACCCAGCTGCTCAAGTCTGTGAGCAGACAGCAGCTGTTAAAG 1196

Db 980 GGAACCAACATTCACCAACCCGTTTGAATATAGTGAACACACTGCAGCTGTTAAAG 1039

Qy 1197 CTTATTCATGCTCTCTCATCTTTCATGGAATCTTTCGATCTCGAGAGGAATCGAGATA 1256

Db 1040 CCATTGCTTGGTCTCTCATGTTTCATGGGCTTCTGCTGCTGGTGGTACTGCTGATA 1099

Qy 1257 GATGATTTGTTTGGATACACACAACTCTACACCTTACCTGATGATGAGCACTGGA 1316

Db 1100 GATGCAATCGTTTGTGGAATACCAACCACTCTATTTAAGTTCCATGATGATTTGCA 1159

Qy 1317 GTCAGGTTTCAATCTTGTCTGTTCCAAAAATGTTCAACGAATAGTAGCAGACATGGGT 1376

Db 1160 GTCAGGTATGCAATCTAGCTTGTCTTAAGAACTGAACAGGCTTGTAGCAGACACCGAT 1219

Qy 1377 ACTCCAGAACACGATTAATTTTGGAGATACCCACATATGTCAAAGTGGCGACTCTTA 1436

Db 1220 ACTCTCAGAACCAATCATTTCTGGAATATACCAACCATGTCCAAATTTCTACTCTAA 1279

Qy 1437 CCGGCATACCTATAGGTTCTCTATCTTGCATCTCTCCAGATGGACAGACTATTGTAA 1496

Db 1280 CCGGTTCACACATACCGAGTCTTATACCTTGGGTTTACCCGATGGACAGAGATTGTA 1339

Qy 1497 CTGAGAGCTGGAGATGAACCGTCTAGGTTCTGGAATGTTTCCCTTCCCTAAATCACA 1556

Db 1340 CAGGAGCAGGAGATGAACCTTAAGTTCTGGAATGTTTCCCTTCCCAAAATCTCAGA 1399

Qy 1557 ATACTGAAGTGAATCGAGCATATCTCTTGAAGAACTACTATCAGGTGA 1609

Db 1400 ACACGATAGTGAATACGGTCTCTTTCTTTGGTAGAACACAACTTCGGTGA 1452

RESULT 4

AY099585

LOCUS

DEFINITION

Arabidopsis thaliana putative fuzzy-related protein (At4g22910)

mRNA, complete cds.

ACCESSION

AY099585

VERSION

AY099585.1

GI:20466238

FLI CDNA.

SOURCE

Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE

1 (bases 1 to 1865)

AUTHORS

Nguyen,M., Karlin-Neumann,G., Southwick,A., Lam,B., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H., Cheuk,R., Chung,M.K., Hayaishizaki,Y., Iehida,J., Kamiya,A., Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K., Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K., Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.

Direct Submission

Submitted (24-APR-2002)

DNA Sequencing and Technology Center,

Stanford University, 855 California Avenue, Palo Alto, CA 94304,

USA

e-mail for correspondence: arab@sequence.stanford.edu

COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the

collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN

Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Iehida,J.,

Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,

Hayaishizaki,Y. and Shinozaki,K.

The Salk, Stanford, RPEC (SSP) Consortium members carried out the

sequencing and annotation of the RAFL cDNAs: Nguyen,M.,

Southwick,A., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J.,

Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung,M.K.,

Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P.,

Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

Nguyen,M. (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed

equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W.

(SSP/Stanford) contributed equally to this work as PIs.

Location/Qualifiers

1. .1865

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="4"

/clone="RAFL09-63-A14 (R24697)"

/notes="This clone is in bluescript

ecotype: Columbia"

1. .1865

/gene="At4g22910"

/notes="synonym: F7H19.90"

67. .1518

gene

CDS

COMMENT

The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RFL CDNA (RAFL CDNA : RIKEN Arabidopsis Full-length cDNA): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, RPEC (SSP) Consortium members constructed and sequenced the PUN1 (ORF) clones using the RAFL CDNA: Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/RPEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/RPEC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

FEATURES
source

Location/Qualifiers

1..1459
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="4"
/clone="U18065"
/note="This clone is in PUN1 51."
ecotype: Columbia

gene

1..1459

CDS

1..1428

/gene="At4g11920"
/codon_start=1

/evidence="experimental"

/product="putative Sw1 protein"

/protein_id="RAFL1474.1"

/db_xref="GI:20259635"

/translation="MEEDESTPPKKSSQNLPPSMNRPTVLESINRLIDSNHYH
SPKPIYDRPTSPKSGNFALFDLASSPNKQKEDGAGSYASLKLALFGPTPE
KSDVNGFSPSNIPFRFTEQRLNLYPPFSDVSVGSPSPKILRSPKYL
DALPQDLYNLVDMSQNLVGLGNCVLMNACSKVKLCLDLGVDETVCSVGLA
LRGTGLALSGSTQVQIWDVLRKNIPTMEHRLRVGLNASSVLSGSSEKILOR
DIRTQEDVSKLKGSSDNRELASGNDNKLFWNQHSITQVLRCEHA
RAVKAIAWSPHFLGASGGTADRCIRFWNTTNHLNCVDNTSQVNLVMSKNVE
LVSTHGYSONQIIVWKYPTMSKLAFLTGHSYRVLVLAIVSPDQTIIVTGAGDETLRFVN
VPPSPKSOSRESEIGALSFGRTIR"

1429..1459

/gene="At4g11920"

393 a 287 c 366 g 413 t

BASE COUNT

ORIGIN

Query Match 35.3%; Score 708.4; DB 8; Length 1459;

Best Local Similarity 71.7%; Pred. No. 2.6e-155;

Matches 957; Conservative 0; Mismatches 371; Indels 6; Gaps 2;

276 TAAGCGGTATGATCAACAGCAACCATACCTCTCTCGAACATCTACTCCGATA 335

101 TCAACCGGTTAATCGATTGGAATCATATCATCTCTCTTCGAAACCTATTACTCAGATA 160

336 GGTTTCATTCGAGTAGATCTGCTCGAAATTCGGTTTGTTCATATCAATATCCGACAG 395

161 GGTTCATTCGAGTAGATCTGCTCGAAATTCGGTTTGTTCATATCAATATCCGACAG 220

396 AAGGACCGGATGATGATTCAGCGCTTATACGATCTTCTGAGAACGGGTTGTTGAC 455

221 CGAATAAGAAAGATGG---GAAAGAGATGGGCTGGTTCTTATCGAGTCTTTTGAAGA 277

456 CGGATGTCGGTCCGGTTACGCGGAAACACGATCGCGCTCGATGACATTCGCGA 515

278 CGGCGCTTTTGGTCGGTGACGCCGAGAAAGATGTTGTTAATGGGTTTTCGGT 337

516 ATAGGAATATTTTAGGTATTAAGACGGAGACGAGACAGTCCATGCATCGCTTCGCGGT 575
338 CGGGGAATATTTTAGGTATTAAGACGGAGACGAGTCTTTGAATTTGATCGCCTT 397
576 TTATCGATGATGATTTTCTCTCGGTGTAAATCATAGTCCGGTTAAGCTCTTAGGAG 635
398 TT---GATTTCTGATGTGGTTAGTGGTGTAGCCCTAGTCTCTTTAAGTCGCGGAGA 454
636 TTCTCTCGATCGCTTATTAAGGTTTGGATGACACCTCTTTGCAAGATGATTTTATCTGA 695
455 TTCTTAGGTCTCTTATTAAGGTTTGGATGAGCGCGCTCTGCAAGATGATTTTACTTGA 514
696 ATCTCGTAGATGCTCTTCAACAAGTGTGGCTGTGGTGTGGTGTGGTGTGGTGTGCTATC 755
515 ATCTCGTAGATGCTCTTCAACAAGTGTGGCTGTGGTGTGGTGTGGTGTGGTGTGCTAT 574
756 TCTCGAATGCTGTAGACGAGGTAACATAATGATGATTTGGGGGTGATGATTTGTTG 815
575 TGTGGATGCTTTGTAGTAGCAGGTAACATAATGATGATTTGGGGGTGATGATTTGTTG 634
816 TTTGTTCTTGTGGTGTGGCTCAACGTGTGATCATCTTTGCTGTGGTGTGGTGTGGTGTG 875
635 TTTGCTCAGTGTGGTGTGGCTCAACGTGTGATCATCTTTGCTGTGGTGTGGTGTGGTGTG 694
876 AAGTTTCAAGTGTGGTGTGGCTCAACGTGTGATCATCTTTGCTGTGGTGTGGTGTGGTGTG 935
695 CAGTACAGATATGGGATGTGTACGATGCAAGATATTAGAAATGAGGTCTATCGGC 754
936 TAGTGTCTGGGCTTTGGCTGTGGTGTGATCATCTTTGCTGTGGTGTGGTGTGGTGTG 995
755 TAAGAGTGTGGCTTTGGCTGTGGTGTGATCATCTTTGCTGTGGTGTGGTGTGGTGTG 814
996 ATATTATCAACGAGATATACGACACAGAGATTTGTTAGTAACTGTGACGACACA 1055
815 GCATACCTTCAGAGAGACATTCGTACTCAAGAGATCATGTCACTAATTAAGGTCA 874
1056 AATCAGAGTGTGTGGACTGAAGTGTGATATGATTAACCGTGTGGTGTGGTGTGGTGTG 1115
875 AATCCGAATCTGTGGACTCAATGCTCTCGACAACTCGCAGCTAGCATCAGGTGGAA 934
1116 ATGACAAATTTGTTTGTGGATCAACATCACTCAACCGCTGTCTCTCAAGTACTGTG 1175
935 ACACAAATAGCTGTTTGTATGAAACCAATTCACCAACCGGTTTGTGAGATCTGTG 994
1176 AGCACACAGCTGTGTTAAGCTATTGCTGTCTCTCTCACTTCACTGACATCTTGTGAT 1235
995 AACACCGCGCAGCAGTGAAGCCATTCGCTGTCTCCATCATCATTTTGGACTTCTTGT 1054
1236 CTGGAGGAGAACTGCAGATAGATGATTGCTGTTTTTGGAAATCAACACAACTCAACC 1295
1055 CTGGTGTGGCACTGCTGATAGATGATTCGCTTTCTGGAACACACACGACAAACACTCAT 1114
1296 TTAGCTGTATGGACACTGGAGTCAAGTGTGCAATCTTGTCTGCTCCAAATATGTCAACG 1355
1115 TAAATTCGTAGATACCAACGACGAGTGTGTAATTTGTTAGTGTCTTGAACGTGATG 1174
1356 AACTAGTAAGCACACATGGGTACTCCAGAACAGATTAATTTGTTGGAGATACCCACTA 1415
1175 AGCTTGTGACACACACGCGGTATTCCAGAACCAAAATCATCGTCTGGAATATCCAA 1234
1416 TGTCAAGCTGGGACTCTTACCGGCATCTTATAGGTTCTCTATCTTGGCATCTCTC 1475
1235 TGTCTAAATTTGGCACTCTCACCTGCTGACCTGACCGCTTCTGACCTTGGGTGTAC 1294
1476 CAGATGACAGACTATTGTAATCTGGAGCTGGAGTGAACCGCTTAGTGTCTGGAATGTTT 1535
1295 CGATGACAGACAAATTTGACAGGAGCAGAGATGAACCTTTGAGATTCGGAATGTCT 1354
1536 TCCCTTCCCTTAATCAAGATATGAAAGTGAATCGAGCATTTATCTCTTGGAGAA 1595
1355 TCCCTTCTCGAAATCTCAGAGCAGGAGCGAAATCGGGGCTATCTTTTGGTAGAA 1414
1596 CTACTATCAGGTGA 1609

Db 2774 CGCGTCGATGACATTCGCGATAGGATATTTTAGGTATAGACGAGACAGACAGT 2833

QY 555 CCATGACATCGCTTTTCGCGTTTATGAGATGATTTTGTTCCTGGTGTAAATCATAGATC 614

Db 2834 CCATGACATCGCTTTTCGCGTTTATGAGATGATTTTGTTCCTGGTGTAAATCATAGATC 2893

QY 615 CGGTAAAGGCTCCTAGGAAGGTTCTCGATCGCCTTATA----- 653

Db 2894 CGGTAAAGGCTCCTAGGAAGGTTCTCGATCGCCTTATAGGTGAATTCATAAAAAAC 2953

QY 654 ----- 653

Db 2954 TTATCTTTTATTTCTTTGATCTTTGTTGAATGTTGTTAGGTTTGTGATTT 3013

QY 654 -----AGTTTTCATGACCTGCTTTGCAAGATGATTTTATCTCAATCTGTAGATT 707

Db 3014 TGATATAGTGTTCGATGACCTGCTTTGCAAGATGATTTTATCTCAATCTGTAGATT 3073

QY 708 GGTCTTCACAAATGTTGGCTGTGTTGGTGAATGTTGTTGTTGTTGTTGTTGTTGTTGTT 767

Db 3074 GGTCTTCACAAATGTTGGCTGTGTTGGTGAATGTTGTTGTTGTTGTTGTTGTTGTTGTT 3133

QY 768 GTAGCAGCAAGTAACTAAATATGTCAT 796

Db 3134 GTAGCAGCAAGTAACTAAATATGTCAT 3162

RESULT 8

BT002165 1511 bp mRNA linear PLN 19-NOV-2002

LOCUS Arabidopsis thaliana cell cycle switch protein (At5g13840) mRNA, complete cds.

DEFINITION

ACCESSION BT002165

VERSION 1

KEYWORDS

SOURCE Arabidopsis thaliana (chale crese)

ORGANISM

REFERENCE

AUTHORS Nguyen,M., Karlin-Neumann,G., Southwick,A., Tripp,M., Miranda,M., Palm,C.J., Bower,L., Jones,T., Banh,J., Carninci,P., Chen,H., Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A., Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K., Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K., Shinzaki,K., Ecker,J., Theologis,A. and Davis,R.W.

TITLE Direct Submission

JOURNAL Submitted (19-NOV-2002) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA

COMMENT e-mail for correspondence: arabesequence.stanford.edu

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Nguyen,M., Southwick,A., Tripp,M., Palm,C.J., Jones,T., Wu,T., Chen,H., Cheuk,R., Chan,M.M., Chang,C.H., Dale,J.M., Deng,J.M., Hsuan,V.W., Lee,J.M., Kim,C.J., Quach,H.L., Shinn,P., Tang,C.C., Toroumi,M., Wallender,E.K., Wong,C., Wu,H.C., Yamada,K., Yu,G., Yuan,S., Ecker,J., Theologis,A. and Davis,R.W.

Nguyen,M., (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W. (SSP/Stanford) contributed equally to this work as PIs.

FEATURES

source

1. .1511

/organism="Arabidopsis thaliana"

gene

CDS

mol_type="mRNA"

db_xref="taxon:3702"

chromosome="5"

clones="U24960"

note="This clone is in pENTR/SD-dTopo. This is a cloned PCR product using RIKEN clone RAFL09-66-B04 (AY099581). as a template"

ecotype: Columbia"

1. .1511

/gene="At5g13840"

1. .1446

/gene="At5g13840"

/codon_start=1

/product="cell cycle switch protein"

/protein_id="AA072176.1"

/db_xref="GI:25084105"

/translation="MASPQSTKTLNLPAGMQTSLRLTFSSFRGISLSLSPKSKT QGSAFPCRSRRHAFDQKEPTTPVREGNEAYRLLKSELFSGDFASPLSPAG TPKHVLDPALDQDFYLVNVDSSNPPSPFSPILGNDGSHSSDSSPPKPRKVPK CSQVMTREGSYISIGTSHGQVQVMDGTQCKRVRTMGHQTGTGLVLAWSRILSSQSRD RNILQHDIRVQSDPVSFKLVHKSSEVCGGLKSHDRELASGNDNOLLVNNHNSQOPI LKLTEHTAAKAITWSPHOSLLASVGSGGTADRCIRFWNTTNGQLNSIDTGSQVQCNLAW SKXNVEIVSTHGYSONQIMLWKYPSKVAITGHSRMLVLYLATSPDGGTITGTGADGE TLRPNVFPSPVQMOTPVKDTGLMSLGRTOIR"

BASE COUNT 411 a 342 c 369 g 389 t

ORIGIN

Query Match 23.5%; Score 472.4; DB 8; Length 1511;

Best Local Similarity 66.1%; Pred. No. 5.7e-100;

Matches 680; Conservative 0; Mismatches 349; Indels 0; Gaps 0;

QY 582 ATGATGATTTTGTTCCTGGTGTAAATCATAGTCGGTAAAGGTCCTAGGAGGTTCTTC 641

Db 419 ATGATAATGTCATCTCAGTGAATCTGCTCCGGCTCCGAAGCCCCCTCCAGGTTCTTA 478

QY 642 GATCCGCTTAAAGGTTTGGATGACACCTGCTTTTGAAGATGATTTTATCTGAATCTGG 701

Db 479 AAACACCTCATAAAGTCTTGATGATGCTCTTCTTTTACAAGATGACTTCTACTTGAATGTTG 538

QY 702 TAGATGCTCTTACACAAATGCTGCTGGTGTGGTGTGGTAACTGCTGCTATCTCTGA 761

Db 539 TGGACTGAGTTTACAGAAATGTTCTTGGCGTTGGGCTTGTACTTGTGTTTATCTTTTGA 598

QY 762 ATGCTTTAGCAGCAAGGTAATTAATATGTAATTTGGGGTTGATGATTTGTTGTTT 821

Db 599 CGGCTTCCAATAGCAAAAGTGAATGATTAATGTGACCTAGGCGCTAATGACAGTGTGCT 658

QY 822 CTGTTGTTGGGCTCAACGTTGCTCATCTTCTGTTGGTAACTAACAATGCTAAAGTTC 881

Db 659 CGGTTCACTGACACAGGAGGTTCTATATATCTTATTTGTTCAAGTCAAGTCAAGTTTC 718

QY 882 AGATTTGGGATGACGCAAGATGCAAGAAGATGAATCAATGGAGGGCCATCGTTTACGTTG 941

Db 719 AGGTTTGGGACGGAAACACATGTCAGAGAGAGTCCGAACCACTGGAGGTCATCAACAAAGAA 778

QY 942 TCGGGGCTTGGGCTCGAGTTTCATCTTTTCTTCTGTTGAGCGGGATAGATATTTT 1001

Db 779 CTGGTGTCTTGGCATGGAATCTCAAGATCTTATCATCAGGAGGACGAGACAGAAATCTCC 838

QY 1002 ATCAACAGATATACGACACAAAGAGATTTTGTGTAGTAACTGTTCAGGACACAAATCAG 1061

Db 839 TTCAAGATGATATCCGAGTCCAGAGCGAATTTTGTTCAGCAAACTCGTGGGCCACAAATCTG 898

QY 1062 AGGTTTGTGACGAGAGTGTGTATATGATAACCGTGTGAGTTGGCATCTGAGGAAATGACA 1121

Db 899 AAGTCTGGGCTCTGAATGTTCTCATGATGATAGAGAGCTTCGGTCTGGAGCAATGATA 958

QY 1122 ACAATGTTGTTTGTGGAATCAACACTCAACCCAGCTGTTCTCTCAAGTACTGTGAGCACA 1181

Db 959 ATCAGTTTATGTTGATGAAACAAATCATTTACAGCAACCACTATTCTGAAGTGTGACTGAGCATA 1018

QY 1182 CAGCAGCTGTAAAGCTATTGTCATGCTCTCTCTCATCTTCTCATGAGTCTTCTTTCATCTGGAG 1241

[illegible]

RESULT 9	AY099581	LOCUS	AY099581	1891 bp	linear	PLN 06-MAY-2002
DEFINITION	Arabidopsis thaliana cell cycle switch protein (At5g13840) mRNA, complete cds.					

COMMENT

USA

e-mail for correspondence: arab@sequence.stanford.edu

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kaniya, A., Sakurai, T., Carninci, P., Kawai, J., Hatawazaki, Y. and Shinozaki, K.

The Salk, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Nguyen, M., Palm, C. J., Southwick, A., Karlin-Neumann, G., Lam, B., Miranda, M., Palm, C. J., Bowser, L., Jones, T., Banh, J., Chen, H., Cheuk, R., Chung, M. K., Kim, C., Lin, J., Liu, S. X., Pham, P. K., Sakano, H., Shinn, P., Yamada, K., Ecker, J., Theologis, A. and Davis, R. W.

(SSP/Stanford) contributed equally to this work as PIs.

FEATURES

source

Location/Qualifiers

1..1891

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="5"

/clone="RAFL09-66-B04 (R24960)"

/note="This clone is in pBluescript

ecotype: Columbia"

1..1891

/gene="At5g13840"

286..1731

/CDS

/gene="At5g13840"

/codon_start=1

/product="cell cycle switch protein"

/protein_id="AAW20433.1"

/db_xref="GI:20466231"

/translation="WASPSQWTKGNLPAQNAQTSRLRLETSSPRGISLSLSSPSKST
 CSDFIPCRSSRLHAFDLDQKEPTTPVREGNEAYSRLLKSELGSPDFASPLLSAPG
 GGGSSAPMSPTCNMLRFTKDRSNSSPSPLGNDLGHSSDSSPPPKPKVPVK
 TPKHLLADPQLDDPYLVNWMSSQNLAVGLGCTVYLTASNKSTKLCDLGGPNDSV
 CSVOTRHSQYTSIGTSHGQVQWMDGTCKRVTGMGHQTRTGVLAWNSRLILSSGSR
 RNILQDRLVQSDPFSYGLKHSEYVCLGWKSHDRSLASGGNDNOLLVNNHNSQOPIIL
 KLTETAAVKAITWSPHSSILIAKSGGTAADRCIRFPWNTNGNLNSIDTGSQVCNLAW
 KUNVEIVSTHSYQNLKRYKPSKVAITLTGHSMRVLYLATSPDGQITVTGAGDE
 TURVNWVPSVKKQTPVKTGUMSLGRQIR"

BASE COUNT 537 a 427 c 422 g 505 t

ORIGIN

Query Match 23.5%; Score 472.4; DB 8; Length 1891;

Best Local Similarity 66.1%; P-red. No. 5.8e-100;

Matches 680; Conservative 0; Mismatches 349; Indels 0; Gaps 0;

QY 582 ATGATGATTTGTTCTGTGTTAATCATAGTCGGGTTAAGGCTCTAGAGAGTTCTTC 641

Db 704 ATGATAATGGTCACTCCAGTGAATCGTCTCCGCTCCGAAGCCCTCGCAAGTTCTCA 763

QY 642 GATCGCTTATAGGTTTGGATGCACCTGCTTTCAGATGATTTTATCTGAATCTGG 701

Db 764 AAACACTCAATAGGTTCTGGATGCTCTCTTTCAAGATGACTTCTACTTGAATGTTG 823

QY 702 TAGATTGGTCTTCACA CAATGTGTTGGCTGTGGTTGGGTAACCTGTGCTATCTCTGGA 761

Db 824 TGGACTGGAGTTCACAGAATGTTCTTTCGCGTTGGGCTTGGTACTTGTGTATCTTTGGA 883

QY 762 ATGCTTTGACGACAGGTAACATAAATATGTCATTTGGGGTTGATGTTGTTGTTGTT 821

Db 884 CGGCTTCCAATAGCAAGTCACTAAGTTAATGACCTAGGGCCCTAATGACAGTGTGCT 943

QY 822 CTGTGTGTTGGGCTCAACCGTGTACTCATCTTCTGTTTGGAACTACAATGTTAAAGTTC 881

Db 944 CGGTTCACTGGACACGGGAAGTTTCATATATATCTATTGGTACAAGTCACTGGTCAAGTTC 1003

QY 882 AGATTTGGGATGAGCAAGATGCAAGAGATAGATCAATGAGGGCCATCGGTTACGTG 941

Db 1004 AGGTTTGGGACGGAAACAAGTGCAGAGAGCTCCGAACCATGGAGGTCACTCAACAGAA 1063

QY 942 TCGGGGCTTTGGCCTGGAGTTCATCTCTTTTGTCTTCTGGTGGACGGGATAAGATATTT 1001

Db 1064 CTGGTGTCTTGGCATGGAATCTCAAGGATCTTATCATCAGGGAGCAGACAGAAACATCC 1123

QY 1002 ATCAACGAGATATACGACA CAAGAGATTTTCTGTAGTAACTGTCAAGGACACAAATCAG 1061

Db 1124 TTCAAGCATGATATCCGAGTCCAGAGCGATTTTGTTCAGCAAACTCTGTGGGCCACAAATCTG 1183

QY 1062 AGGTTTGTGACATGAAGTGTTCATATGATAACCGTGAAGTTGGCATCTGGAGGAAATGACA 1121

Db 1184 AAGTCTCGGCTCGAATGGTCTCATGATGATAGAGAGCTTGGCTCTGGAGGCAATGATA 1243

QY 1122 ACAAAATGTTTGTGGAATCAACATCAACCCAGCTGTCTCTCAAGTACTGTGAGGACACA 1181

Db 1244 ATCAGTATTGTTATGAAACAATCATTTCAAGCAACCTATTTCAAGCTGACTGAGCATA 1303

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAX plate: 96 Row: k Column: 23
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis. Similarity but not identity to protein.

	FEATURES	SOURCE
1.	100% Cotton	100% Cotton
2.	100% Polyester	100% Polyester
3.	100% Nylon	100% Nylon
4.	100% Rayon	100% Rayon
5.	100% Linen	100% Linen
6.	100% Silk	100% Silk
7.	100% Wool	100% Wool
8.	100% Leather	100% Leather
9.	100% Rubber	100% Rubber
10.	100% Glass	100% Glass
11.	100% Metal	100% Metal
12.	100% Plastic	100% Plastic
13.	100% Paper	100% Paper
14.	100% Fabric	100% Fabric
15.	100% Wood	100% Wood
16.	100% Stone	100% Stone
17.	100% Concrete	100% Concrete
18.	100% Brick	100% Brick
19.	100% Tile	100% Tile
20.	100% Marble	100% Marble
21.	100% Granite	100% Granite
22.	100% Quartz	100% Quartz
23.	100% Sapphire	100% Sapphire
24.	100% Emerald	100% Emerald
25.	100% Ruby	100% Ruby
26.	100% Diamond	100% Diamond
27.	100% Gold	100% Gold
28.	100% Silver	100% Silver
29.	100% Platinum	100% Platinum
30.	100% Palladium	100% Palladium
31.	100% Rhodium	100% Rhodium
32.	100% Iridium	100% Iridium
33.	100% Osmium	100% Osmium
34.	100% Rhenium	100% Rhenium
35.	100% Ruthenium	100% Ruthenium
36.	100% Cobalt	100% Cobalt
37.	100% Nickel	100% Nickel
38.	100% Copper	100% Copper
39.	100% Zinc	100% Zinc
40.	100% Aluminum	100% Aluminum
41.	100% Magnesium	100% Magnesium
42.	100% Titanium	100% Titanium
43.	100% Vanadium	100% Vanadium
44.	100% Chromium	100% Chromium
45.	100% Manganese	100% Manganese
46.	100% Iron	100% Iron
47.	100% Steel	100% Steel
48.	100% Stainless Steel	100% Stainless Steel
49.	100% Inconel	100% Inconel
50.	100% Monel	100% Monel
51.	100% Hastelloy	100% Hastelloy
52.	100% Superalloy	100% Superalloy
53.	100% Composites	100% Composites
54.	100% Carbon Fiber	100% Carbon Fiber
55.	100% Kevlar	100% Kevlar
56.	100% Fiberglass	100% Fiberglass
57.	100% Resin	100% Resin
58.	100% Epoxy	100% Epoxy
59.	100% Polyurethane	100% Polyurethane
60.	100% Acrylic	100% Acrylic
61.	100% PVC	100% PVC
62.	100% PE	100% PE
63.	100% PP	100% PP
64.	100% PS	100% PS
65.	100% PMMA	100% PMMA
66.	100% PC	100% PC
67.	100% ABS	100% ABS
68.	100% PA	100% PA
69.	100% PET	100% PET
70.	100% PBT	100% PBT
71.	100% POM	100% POM
72.	100% PPS	100% PPS
73.	100% PI	100% PI
74.	100% PVDF	100% PVDF
75.	100% PTFE	100% PTFE
76.	100% FEP	100% FEP
77.	100% EPE	100% EPE
78.	100% EVA	100% EVA
79.	100% PU	100% PU
80.	100% TPU	100% TPU
81.	100% SBR	100% SBR
82.	100% NBR	100% NBR
83.	100% CR	100% CR
84.	100% EPDM	100% EPDM
85.	100% CSM	100% CSM
86.	100% SI	100% SI
87.	100% BR	100% BR
88.	100% NR	100% NR
89.	100% SBR	100% SBR
90.	100% NBR	100% NBR
91.	100% CR	100% CR
92.	100% EPDM	100% EPDM
93.	100% CSM	100% CSM
94.	100% SI	100% SI
95.	100% BR	100% BR
96.	100% NR	100% NR
97.	100% SBR	100% SBR
98.	100% NBR	100% NBR
99.	100% CR	100% CR
100.	100% EPDM	100% EPDM

```

1. .3033
  /organism="Xenopus laevis"
  /mol_type="mRNA"
  /db_xref="taxon:81355"
  /clone_MGC:53235 IMAGE:554
  /tissue_type="Embryo, stage"
  /clone_lib="NICHG_XGC_Embd"
  /lab_host="DH10B"
  /note=/vector: pCMV-SPORT6
161. 634

```

50

BASE COUNT	ORIGIN
842 a	615 c 659 g 917 t

Query Match 20.2%; Score 405.2; DB 5; Length 3033;
Best Local Similarity 66.2%; Pred. No. 3.3e-84;
Matches 616; Conservative 0; Mismatches 308; Indels 6; Caps 2;

Qy	613	TCGGTTAAGGCTCCTTAGGAAGGTTTCCTCGATCGCCCTTATAAGGTTTTGGATGCGACCTGC	672
Db	736	TCCACGAAAGCCTACAAGGAAGATTTCCGAAGATCCCTTTAAAGTTCTGATGCACTGA	795
Qy	673	TTTCSAAGATGATTTTTTATCTGAAATCTGCTAGATTTGGTCTTTCACAAATGTGTGGCTGT	732
Db	796	GTTACAAGATGACTTTTACTCTAACCTAGTGGATTTGGTCACTCCCTCAATGTTTTAAGTGT	855
Qy	733	TGGTTTCGGTAACTGTGCTCATCTCTGGAAATGCTTTGTAGCAGCAAGGTAACTAAATTATG	792
Db	856	GGGCTTGGGAACCTTGTGTGTACTATTGCACTGCATGTACTAGCCAGGTGACTAGACTATG	915
Qy	793	TGA---TTTGGGGGTTGATGATGTGTTTCTTCTGTTTGGTTGGCTCAAAGTGTACTCA	849
Db	916	TGACCTCTCTGTGGAAGGGGATTCGTAACTTCAGTTGGTTGGTCAGAGCGGGAAACCT	975
Qy	850	TCTTGCTGTTTGGAACTAACAAATGGTAAAGTTTCAGATTTGGGATGCAGCAGAGATCAAGAA	909
Db	976	GGTAGCTGTGGGGACACACAAGGCTTTGTTTCAGATTTGGATGTCATCAGCAGGGAAGAA	1035
Qy	910	GATAAGATCAATGAGGGGCCATCCGGTTACGTTGTCGGGCCCTTGGCCTGGAGTTCACTCT	969
Db	1036	ACTGTCCACTCTGAGGGGACACACAGCCCGTGTGGGAGCCCTTGGCTTGGAAATGCAGACCA	1095
Qy	970	TTTCTCTCTCGTGGAGGGGATAAGATATTTATCAACGAGATATGSCACAA---GA	1026
Db	1096	GCTTTCTCTCGGGATGGGACAGATGATCTTCAAAAGGATTAATCGCACACCACTGT	1155
Qy	1027	AGATTTGTTTAACTGTTCAGGACACAAATCAGAGTTTGTGCACTGAAGTGGTCATA	1086
Db	1156	TCAGTCAAGAGAAGGCTGCAAGGCCACAGGCAAGAGTGTGTGGCCTCAAGTGGTCCAC	1215
Qy	1087	TGATAACCGTAGTTGGCATCTGGAGGNAATGACAAATAATTGTTTGTTTGGATCAACA	1148
Db	1216	AGACCACCAGCTCCTGGGCTTCAGAGGGAATGACAAAGGCTCTTGGTTTGGAAACCATTC	1275
Qy	1147	CTCAACCCAGCCTGTTCCTCAAGTACTGTGAGCACACAGCAGCTGTATAAAGCTATTGGCATG	1206
Db	1276	TAGCCTGAGTCTCTGTTCAAACAGTATATCTGAACATCTAGCTGCAGTGAAGGCAATGGCCTG	1333
Qy	1207	GTCTCCTCATCTTTCATGCACTTCTTGCACTCTGGAGGGAAGTCTGCAGATAGATGATATTCG	1266

clone distribution: MGC clone distribution information can be found

Db 1336 GTCCCTCATCAACATGGCTTTTGGCTTCAGGCGGGGCACTGCTGACCGCTGCATCCG 1395
QY 1267 TTTTGGATACACACAACTCACACCTTAGCTGTATGGACACTGGAAGTCAGGTTTG 1326
Db 1396 CTTCTGGAACACATACAGGCGAGCTTTTACAGTGTATGTATGATCTGGCTCTCAAGTCG 1455
QY 1327 CAATCTTGTCTGTGTCACAAAATGTCACGAACCTAGTAAGCACACATAGGTACTCCAGAA 1386
Db 1456 TAACTTGGCTGTCTAAGCATGCAATGAATCAATGATGGTAAGCACACATGGCTACTCTCAAA 1515
QY 1387 CAGATATTGTTGAGATACCCCACTATGTCACAAAGCTGGGCACTCTTACCGGCATAC 1446
Db 1516 TCAGATATTAGTATGAAATATCCATCGCTTACACAGGTGGCCAGATTGACTGGACATTC 1575
QY 1447 TTATAGGTTCTCTATCTTCCATCTCTCCAGATGACACAGATATTGTAACCTGGAGCTGG 1506
Db 1576 CTACCGTGTCTCTACTCGCTATGTCACAGATGGGAGGCCATTGTTACAGGAGCAGG 1635
QY 1507 AGATGAACCGCTAGGTTCTGGAATGTTTT 1536
Db 1636 TGATGAACGCTGAGATTCTGGAATGTTTT 1665

RESULT 11

XLNFRP 2649 bp mRNA linear VRT 08-SEP-1997
LOCUS Xenopus laevis mRNA for Fizzy-related protein.
DEFINITION
ACCESSION Y14163
VERSION Y14163.1 GI:2326942
KEYWORDS fizzy-related protein; fzr gene.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Fipidae;
Xenopodinae; Xenopus.

REFERENCE

1 Sigrist S.J. and Lehner, C.F.
Drosophila fizzy-related down-regulates mitotic cyclins and is
required for cell proliferation arrest and entry into endocycles
Cell 90 (4), 671-681 (1997)

JOURNAL

MEDLINE
97433078
PUBMED
9288747

REFERENCE

2 (bases 1 to 2649)

AUTHORS

Lehner, C.F.

TITLE

Direct Submission

JOURNAL

Submitted (04-JUL-1997) C.F. Lehner, Department of Genetics,
University of Bayreuth, 95440 Bayreuth, FRG

FEATURES

Location/Qualifiers
1..2649
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/tissue_type="ovary"

gene

1..2649

5'UTR

10..99

CDS

100..1581

/gene="fzr"
/codon_start=1
/product="fizzy-related protein"
/protein_id="CAA74576.1"
/db_xref="GI:2326943"
/db_xref="SPTRMBL:O42585"
/translation="MDQYERLLRQINLQNETIPICASEMRETLTPNPSPSKH
GDRTPRAGANVSNINFRINENKSPONRKADATADSGKGLAYSALLKNEILGA
GIEKVPQEDERLQSPTEKSLFTYSLSKRAPDDGNEVSPYSLSPVSNKSKL
LRSRPRTRIKSPIKPKVLDAPELODPYLVNWSLAVLVSVGLACTCYVLWSACTSQ
VTRCLDSVSGDSVTSVGNBSRGNLVAVGTHKGFVQIWDASAKKSLSTEGHTARVGA
LAWNDQLSSGSRDMTLQDRTFPVQSERLRQHRQEVCGKLVSTDDQLLASGND
NKLIVNWSLSLSPVQYTHEAAVKALAWSPHQGLASGGTADRCIRFWNTLQGP
LOCTDTSQVNCNLAWSKANELVSTHGYSQNLVWKYPSLTQVAKLTGHSRVLYLA
MSPDGEALVTGABETLRFWNVFSKTRSTKESVNLNLFTRIR"


```

RESULT 12
BC048038      2089 bp      mRNA      linear      VRT 05-MAR-2003
LOCUS         Danio rerio, clone MGC:55450 IMAGE:2639752, mRNA, complete cds.
DEFINITION    BC048038
ACCESSION     BC048038
VERSION       BC048038.1 GI:28856199
KEYWORDS      MGC.
SOURCE        Danio rerio (zebrafish)
ORGANISM      Danio rerio
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
              Cypriniformes; Cyprinidae; Danio.
REFERENCE     1 (bases 1 to 2089)
AUTHORS       Strausberg, R.
TITLE         Direct Submission
JOURNAL       Submitted (04-MAR-2003) National Institutes of Health, Mammalian
              Gene Collection (MGC), Cancer Genomics Office, National Cancer
              Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
              USA
REMARK        NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT       Contact: MGC help desk
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: Dr. Sumio Sugano and Dr. Ko-ichi Kawakami
              cDNA Library Preparation: Dr. Sumio Sugano and Dr. Ko-ichi Kawakami
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Institute for Systems Biology
              http://www.systemsbio.org
              contact: amadan@systemsbio.org
              Anup Madan, Jessica Fahay, Erin Helton, Mark Kettelman, Anuradha
              Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAX Plate: 101 Row: h Column: 12.
              Location/Qualifiers
                1..2089
                  /organism="Danio rerio"
                  /mol_type="mRNA"
                  /strain="AB"
                  /db_xref="taxon:7955"
                  /clone="MGC:55450 IMAGE:2639752"
                  /tissue type="Whole body, adult, (one male and one female,
                  including unfertilized eggs)"
                  /clone_lib="Sugano Kawakami zebrafish DRA"
                  /lab_host="DH10B"
                  /note="Vector: pME18S-FL3"
                  227..1714
                    /codon_start=1
                    /product="Unknown (protein for MGC:55450)"
                    /protein_id="AAH48038.1"
                    /db_xref="GI:28856200"
                    /translation="WDQYERLLRQINVDNTSPMKADVILCHLTPTPEPLSGPS
                    KHGRFIPFRAGAWNSINFRINENKSPSNKTKDSDSKADGLAYSALIKNEL
                    LGAGLEKVDPTEDRRLOQPTPERRSLFYSLSAKRSTDDNSVSPISLPSVSSQ
                    KLRSPRPKIKIPKVLDAPELODDFYLNLDWSLNVLVSLVIGLTCVILWACT
                    SQVRLCDLSVEGDSVTSGVSRGNLAVGTHGFVQIWDATAGKKLFALBHTARV
                    GALANWADLSSGRDRMLQDRTPLOSERLROHROEVCKLWSTHQLIASGG
                    NDNLVLWNHSSVLPMQYTERHLAAKALAWSPHOGLLASGGGTADRCIRFWNTLTA
                    OPLCIDTGSQVCLNLSKTNELVSTHGYSONQLVWKYPSLTQVAKLAGHSRVLY
                    LAMPFDGRAIVTGAGDETLRFNWFPSKTRSTKESVSVINLFTIR"
BASE COUNT   599 a 495 c 467 g 528 t
ORIGIN
Query Match      17.8%; Score 357; DB 5; Length 2089;
Best Local Similarity 63.4%; Pred. No. 6.7e-73;
Matches 581; Conservative 0; Mismatches 330; Indels 6; Gaps 2;

QY 647 CTTTAAAGTTTGGATGCACTCTTTGCAAGATGATTTTATCTGAATCTCGTAGAT 706
DB 755 CCGTTAAAGTTCITGATGACACAGAGCTACAGGACGACTTCTACCTCAACTTAGTGAC.814
QY 707 TGGTCTTCACAAATGTGTGGCTGTGTTGGTAACTGTGTATCTCTGGAATGCT 766

```

```

Db 815 TGGTCTCTTTAAATGTTCTCAGTGTGGGTCTGGAACAATGTGTTTATCTTATGAGTGCC 874
QY 767 TGTAGCAGCAGGTAACTAAATTAATGTA---TTTGGGGTTGATGATGTTGTTTCT 823
DB 875 TGCACCACTCAAGTGAACGGCTATGTGACTTAATGAGGGCGATTCGTCACTCT 934
QY 824 GTTGGTTGGGCTCAACGTGTPACTCATCTTCTCTTGGAACTAACAATGGTAAAGTTGAG 883
DB 935 GTGGGCTGCTCAGAGAGGGGAAATCTTGTCTGCTGGGAAACACACAAAGGCTTTGTACAA 994
QY 884 ATTTGGGATCAGCAAGATGCAAGAAGATAAGATCAATGGAGGCCATCGGTACGTGTC 943
DB 995 ATTTGGGATGCCCGGTGGCAAAAACCTGTTGCTTGGAGGACACACTGCAAGATT 1054
QY 944 GGGGCTTGGGCTGGAGTTTCATCTCTTTTGTCTTCTGTTGACGGGTAAGAATATTTAT 1003
DB 1055 GGTGCAATTAGCATGGAATGCAGATCAGTTGTCAACAGCAGTCTGTGACAGATGATCTC 1114
QY 1004 CAACAGATATACGCA---CACAGAAGATTTTGTAGTAACTGTCAAGGACACAAATCA 1060
DB 1115 CAGAGAGATATCAGGACTCTCTCTCTGCAAGCGCAGGCTGCGGGCCACAGACAG 1174
QY 1061 GAGGTTTCTGAGTCAAGTGTTCATATATCAATACCGTGAAGTTGGCATCTGGAGGAATGAC 1120
DB 1175 GAGGTTCTGGGACTCAAGTGGAGCAGTCAACCATCAGCTGCTCGCTCTGGGGAAATGAC 1234
QY 1121 AACAAATGTTTGTGGAAATCAACATCAACCCAGCTGTCTTCAAGTACTGTGAGCAC 1180
DB 1235 AATAAGCTGTTAGTGTGGAATCACTCAAGTGTGTACCAATGCAGCAGTCACTGAAAC 1294
QY 1181 ACAGCAGCTGTTPAAAGCTATTGATGTTCTCTCATCTTTCATGGACTTCTTGCACTCGGA 1240
DB 1295 TTGGCAGCAGTGAAGCCCATCGCTGGTCTCTCTCATCAACACGGGTACTGGCGTCTGA 1354
QY 1241 CGAGAACTGCAAGATGATGATTTGCTTTTGGAAATACAAACCAAACTCACACCTTACG 1300
DB 1355 GGGGGAACCGGCGACCGCTGTATTGCTTTTGGAACTCTCACAGCAGCAGCCTGCGAG 1414
QY 1301 TGTATGGACACTGGAAGTCAAGTTTGGCAATCTTGTCTGTCGTCGTCGTCGTCGTCGTCG 1360
DB 1415 TGCATCGACACCGGCTCTCAAGTCTGCAATCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1474
QY 1361 GTAAGCACACATGGGTACTCCAGAACAGATATTGTTGGAGATACCCCACTATGTCA 1420
DB 1475 GTCAGCACACACGGTTACTCTCAGAATCAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1534
QY 1421 AAGCTGGGACTTTTACCGGCCATATCTATAGGTTCTCTATCTTGGCATCTCTCCAGAT 1480
DB 1535 CAAGTAGCCAAACTGACCGGCCACTCTTATAGGGTGTCTTATCTGGCAATGTCTCCAGAT 1594
QY 1481 GGACAGACTATTGTAACTGGAGCTGGAGATGAACGCTTAGGTTCTGGAATGTTTTCCT 1540
DB 1595 GGAGAGCGGATAGTACCGGAGCAGCGATGAATGAACTCTGCGCTTTTGGAAATGTTATTAGT 1654
QY 1541 TCCCTTAATCAGAA 1557
DB 1655 AAAACACGATCAACAAA 1671

RESULT 13
BC006616      1861 bp      mRNA      linear      ROD 16-APR-2003
LOCUS         Mus musculus frizzy-related protein 1 (Drosophila), mRNA (cdna
DEFINITION    clone MGC:11723 IMAGE:3967009), complete cds.
ACCESSION     BC006616
VERSION       BC006616.1 GI:13879283
KEYWORDS      MGC.
SOURCE        Mus musculus (house mouse)
ORGANISM      Mus musculus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE     1 (bases 1 to 1861)

```


AUTHORS

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L., Schaefer, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Frange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A.C., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalley, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2238257

12477932

2 (bases 1 to 1861)

Strausberg, R.

Direct Submission

Submitted (27-APR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs-remail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amg@bcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louissege, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 17 Row: 1 Column: 12

This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

FEATURES

Location/Qualifiers

1..1861

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="MGC:11723 IMAGE:3967009"

/tissue_type="Mammary tumor. Metallothionien-TGP alpha model. 10 month old virgin mouse. Taken by biopsy."

/clone_lib="NCI CGAP Mam1"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

1..1861

/gene="Pzrl1"

/note="synonym: Pzr"

/db_xref="LocusID:56371"

/db_xref="MGI:1926790"

224..1705

/codon_start=1

/product="Pzrl protein"

/protein_id="AAH06616.1"

/db_xref="GI:13879284"

gene

CDS

/db_xref="LocusID:56371"

/translation="MPDVERLLRRIILIONENTVPCVSEMRRLTPANSPVSSPSKHH
GDRFIFRAGANWVNFHINENEKSPQNRKAKDATSDNGKGLAYSALLKNEILGA
GIEKQDPKTRIKIPFKVLDPAELQDFYLNLDVSNLVSGLGTCVYWSAKTSQ
LRFPKPTKIKIPFKVLDPAELQDFYLNLDVSNLVSGLGTCVYWSAKTSQ
VTRLCDSVEGSDSVTSVGNRNLVAVGTHKGFVQVWDAAGCKGSLMEGHFARVGA
LAWNADQLSSGSDRMILQDRIPTPLQSRRLQHQVQKGLKSTHDLGLASGND
NKLNVNHSLSLPVOOYTHELAAVAIAWSPHOGLLASGGGTADRCIRFWNTLTGP
LCQIDTGSOVNLSKSHNELVSTHYSONQILVMKYPSLTQVAKLTGHSYRVLYLA
MSPDGEALVTGAGDETLRFNVPVSKTSKESVSVLNLFIR"

BASE COUNT 409 a 560 c 562 g 330 t

ORIGIN

Query Match 17.6%; Score 352.8; DB 10; Length 1861;
Best Local Similarity 62.9%; Pred. No. 6.4e-72;
Matches 581; Conservative 0; Mismatches 337; Indels 6; Gaps 2;

QY 620 AAGCTCTAGGAAGTTCTCGATCGCCTTAAAGGTTTGGATGACACCTGTTGCAA 679
DB 719 AAGCCACACGCAAGATCTCTAAGATTCCCTTCAAGGTCTGGACGGCCAGAGCTTCAG 778
QY 680 GATGATTTTATCTGAATCTGGTAGATTGGTCTTTCACACAATGTGTGGCTGTGGTTTG 739
DB 779 GACGACTTCTACCTCAACTTGGTGGATGGTCTCTCCCTCAACGTCTCAGCGTGGGCTG 838
QY 740 GGTAACTGTGTCTATCTCTGGAATCTGTAGAGCAAGTAATTAATTTATGATTTG 799
DB 839 GGCACCTCGGTGTACCTGTGTGAGTGCATGCACACGACAGGTGACCCGGCTCTGTGACCTC 898
QY 800 GGGGTGTA---TGATTGTGTGTCTGTGGTGGGCTCAACGTGTGTCTCATCTTTGCT 856
DB 899 TCTGTAGAGGGGACTCAGTGACTTCCGTGGTGGTCTGAGCGGGGAAGTTGGTGC 958
QY 857 GTTGAACCTAACATGTGAAGTTTCAATTGGGATGAGCAAGATGCAAGATGAAGA 916
DB 959 GTAGTACACACAAGGGCTTCTGTGCAGATCTGGGACGCTGTCTGTGGGAAGAGCTGTCC 1018
QY 917 TCAATGAGGGCCCATCGGTACGTTACGTTGGGCGCTTGGCTGAGTTCATCTCTTTGCT 976
DB 1019 ATGCTGGAGGGCCACACAGCAGAGTGGGGCGCTGGCTGGATGTGACCACTGTGCA 1078
QY 977 TCTGTGACGGGATAGAATATTATCAACGAGATATACGACACAA---GAAGATTTT 1033
DB 1079 TCTGTAGCCGTGACCGCATGATCTACAGGGGATATCCGACACACACCCCTGCAGTCA 1138
QY 1034 GTTAGTAACCTGTACGACACAAATCAGAGTTTGTGGACTGAGTGGTCATATGATAAC 1093
DB 1139 GAGCGCGGCTGACAGGGCCACCGGAGGAAGTGTGGCCCTAAAGTGGTCACAGACCAC 1198
QY 1094 CGTGAGTTGGCATCTGGAGGAAATGACAACTTTGTTTGTGGAAATCAACACTCAACC 1153
DB 1199 CAGCTGTGCTCGGCGGGCAATGACAACTGCTGCTGTGGNAACCTCTAGTCTA 1258
QY 1154 CAGCTGTGCTCCTCAAGTACTGTGAGGACACAGCAGTGTAAAGTATTGCACTGCTCT 1213
DB 1259 AGCCCTGTGCAGCAGTATACGGAGCACCTGACGCGGTGAAGGGCTATTGCGTGGTCCC 1318
QY 1214 CATCTTCATGACTTCTTTGCACTTGGAGGAGAACTGCAGATAGATGATTTCGTTTTGG 1273
DB 1319 CACACAGATGACTGCTGTGCATCTGGTGTGGCAGGCTGACCGCTGTATCCGATTCTGG 1378
QY 1274 AATACACACCAAACTCAACCTTAGCTGTATGGACACTGGAAGTCAGGTTTGCATCTTT 1333
DB 1379 AACACTCTGACAGGTGACGCCACTGCAGTGCATTGACACAGGCTCAAGAGTGTGCAACCTG 1438
QY 1334 GTCTGTTCCAAATAATGTCACGAAGTACTAGACACATGGGTACTCCAGAACAGATTT 1393
DB 1439 GCCTGTTCCAGACAGCCCAATGAGCTGTGTGAGCACAATGTGCTACTCACAGAACAGATC 1498
QY 1394 ATTGTTGGAGATACCCCACTATGTCAAAGCTGGGAGCTTTACCGGCCATCTTATAGG 1453
DB 1499 CTCGTGTGGAAGTACCCGCTCCCTTACGCGAGGTGGCCAGCTCACTGGCCACTCATATCGT 1558

Qy 1454 GTTCTCTATCTCCCATCTCTCCAGATGGACAGACTATTGTAACTGAGCTGAGATGAA 1513
Db 1559 GTCTCTACCTGCTGCTCCCTGATGGGAGGAGCCATAGTACCGGAGCTGAGATGAG 1618
Qy 1514 ACCTTAGGTTCTGGATGTTTTC 1537
Db 1619 ACCTGAGGTTCTGGATGTTTTC 1642

RESULT 14

AF083809 2258 bp mRNA linear ROD 01-SEP-1999
DEFINITION Mus musculus fizzy-related protein (Fyr) mRNA, complete cds.
ACCESSION AF083809
VERSION AF083809.1 GI:5813824
KEYWORDS
SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

Jin, D.-Y. and Jeang, K.-T.
1 (bases 1 to 2258)
Characterization of mouse fizzy-related protein

JOURNAL

Unpublished
2 (bases 1 to 2258)

AUTHORS

Jin, D.-Y.

TITLE

Direct Submission

JOURNAL

Submitted (12-AUG-1998) LMM, NIAID/NIH, 9000 Rockville Pike,
Bethesda, MD 20892-0460, USA

FEATURES

Location/Qualifiers
1..2258

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

1..2258

/gene="Fyr"

143..1624

/gene="Fyr"

/function="cell cycle regulator"

/codon_start=1

/product="fizzy-related protein"

/protein_id="AA052029.1"

/db_xref="GI:5813825"

/translation="MDQYERLLRQIIQNTVPCVSEMRLLTPANSPVSPSKH
GDRFIPRAGANWVFNHRIENKSPQNRKADATDNGKDLGAYALLKNEELGA
GIERKVPQEDRLQPTPEHKLFTYLSKRSPPDNDVSPVSLVSNKSKL
LRSRKPRKTSKIPKVLDAPELDQDFYLNVDWLSNLVSLGATCYVLWSACTSQ
VTRCDLSVEGDSVTSVGSERGNLVAVGTHKGFQIWDAAAGKLSMLEHTARVGA
LAWADQLSSSRDRMILORDIRTPLOSERLQHRQEVCKLWSDHOLLASGND
NKLIVNHNSSSPVQYIYELHAAKALAWSPHQHLLASGGGTADRCIRFWNTLTGP
LOCIDTSGVQCNLAWSKHANLSTHGTSTQNLIVWKYPSLTQVAKLTGHSYRVLYA
MSPDGEAIVTGAGDETLRFWNVFKSTKESVSLNLFTRIR"

BASE COUNT 463 a 702 c 656 g 437 t

ORIGIN

Query Match 17.5%; Score 351.2; DB 10; Length 2258;
Best Local Similarity 62.8%; Fred. No. 1.5e-71;
Matches 580; Conservative 0; Mismatches 338; Indels 6; Gaps 2;
Qy 620 AAGGCTCTAGGAAGTTCTCGATCGCCTTATAGGTTTGGATGCACTGCTTGCAA 679
Db 638 AAGCCACAGCAAGATCTTAAGATTCCTTCAGGTGCTGGACGCGCAGAGCTTCAG 697

Qy

680 GATGATTTTATCTGAATCTGTAGATGGTCTTCACACATGTTGGCTGTTGTTG 739

Db

698 GACGACTTCTACTCAACTGTGTGACTGGTCTCCCTCAACGTCTCAGCGTGGGGCTG 757

Qy

740 GGTAACTGTGTCTACTCTGGAATGCTGTAGCAGAGGTAACATAATTATGATTG 799

Db

758 GGCACCTGCTGCTACTGTGGAGTGCATGCACACGAGGTGACCGGCTCTGTGACCTC 817

Qy

800 GGGGTGA---TGATGTGTTTGTCTGTGTTGGCTCAACGTTGACTCATCTGCT 856

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

Db 818 TCTGTAGAGGGGACTCAGTGACTTCGTTGGCTGGTCTGAGCGGGGAACTTGGTCGCA 877
Qy 857 GTTGAACCTAACATAGTAAAGTTTCAGATTGGATGACAGACATGTCAGAGATGAAGA 916
Db 878 GTAGGTACACAAAGGGCTTCGTGCAGATCTGGGACGCTGCTGCTGGGAGAGACTGTCC 937
Qy 917 TCAATGGAGGGCCATCGTTACGTGTCGGGCGCTTGGCCCTGGAGTTTCATCTCTTTTGTCT 976
Db 938 ATGCTGGAGGGCCACACAGACAGAGTGGGGCGCTGGCTGGATGCTGACCACTGTCTCA 997
Qy 977 TCTGGTGAACGGGATAGAATATTATCAACGAGATATATGACACAA---GAAGATTTT 1033
Db 998 TCTGTAGTCCGCTGACCGCATGATCTACAGCGGATATCCGACACACCCCTCGATGCA 1057
Qy 1034 GTTAGTAACTGTGAGGACACAAATCAAGAGTTTGTGACTGGAAGTGGTCTATATGATAAC 1093
Db 1058 GAGCGCGGCTGACGGCCACCGCAGAGAGTGTGTGGCTTAAAGTGGTCCACAGACCAC 1117
Qy 1094 CGTAGTTGGCATCTGAGAGGAAATGACAAATAATTTTGTGGATGCAACACTCAACC 1153
Db 1118 CAGCTGCTTGGCTCGGGGGCAATGACAAAGCTGCTCGTGTGGAAACCACTCTAGTCTA 1177
Qy 1154 CAGCTGTCCTCAAGTACTGTGACACACAGCAGCTGTTAAAGCTATTGCAATGCTCTCCT 1213
Db 1178 AGCCCTGTGCAGAGTATACGAGACACTGCGACCGCTGAGGCTATTGCTGCTCCCA 1237
Qy 1214 CATCTTCACTGACTTCTTGCATCTGAGAGGAACTGACAGATAGATGATTGCTTTTGG 1273
Db 1238 CACCAGCATGACTGCTGGCATCTGGTGGCACGGCTACCGCTGCATCCGATTTCTGG 1297
Qy 1274 AATCAACCAACAACTCAACCTTAGCTGATGACACTGGAAGTCAAGTTTGCATCTT 1333
Db 1298 AACACTGTGACAGGTGACGCCACTGCACTGCATGACACAGGCTCACAGTGTGCACCTG 1357
Qy 1334 GTCTGCTCCAAAATGTCAAGCACTAGTAAGCACACATGGGTACTCTCCAGAACAGATT 1393
Db 1358 GCCTGCTCCAGCAGCCCAATGAGTGTGTGAGCACACATGGCTACTCACAGAACCCAGATC 1417
Qy 1394 ATTGTTGGAGTACCCCACTATGTCAAGCTGGCGCTCTTACCGGCCATCTTATAGG 1453
Db 1418 CTCGTGTGGAAGTACCCCTGCTTACGAGGTGGCAGCTCACTGCGCATCTCATATGCT 1477
Qy 1454 GTTCTCTATCTTCCCATCTCTCCAGATGACAGACTATTGTAATCTGAGCTGGAGATGAA 1513
Db 1478 GTCTCTACTGCGCATGCTCCCTGATGGGAGGCCATAGTACCGGAGCTGGAGATGAG 1537
Qy 1514 AGCTTAGGTTCTGGAATGTTTTC 1537
Db 1538 ACCCTGAGGTTCTGGAATGTTTTC 1561

RESULT 15

AB013462

LOCUS

DEFINITION

AB013462

ACCESSION

VERSION

AB013462.1

GI:6463678

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1

Kotani, S., Oyama, T. and Todokoro, K.

Human homologue of Fizzy-related protein

Published only in Database (1999)

2 (bases 1 to 1491)

Kotani, S., Oyama, T. and Todokoro, K.

Direct Submission

Submitted (29-APR-1998) Kazuo Todokoro, The Institute of Physical

and Chemical Research (RIKEN), Molecular Cell Science Laboratory;

3-1-1, Koyadai, Tsukuba, Ibaraki 305-0074, Japan

(E-mail: todokoro@rtc.riken.go.jp, Tel: 81-298-36-9075,

Db 1123 GGCGGCACAGCTGACCGCTGTATCGCTTCTGGAACACGCTGACAGGACCAACCACTGCAG 1182
Qy 1301 TGTATGGACACTGGAGTCAAGCTTTTGGCAATCTTGTCTGTCACAAAATGTCAACGAACCTA 1360
Db 1183 TGTATCGACACCGGCTCCCAAGTGTGCAATCTGGCTGTGTCAGCAACGCCAACGAGCTG 1242
Qy 1361 GTAAAGCACACATGGGTACTCCAGAACCAAGATTATTGTTTGGAGATACCCCACTATGTCA 1420
Db 1243 GTGAGCACGACGCTTACTTACAGAACCAAGATCCTTGTCTGGAAGTACCCCTCCCTGACC 1302
Qy 1421 AGCTGGGACCTTTACCGGCCATCTTATAGGTTCTTATCTTCTGTCATCTTCCAGAT 1480
Db 1303 CAGGTGGCCCAAGCTGACCGGGCACTCTTACCGCGTGTCTTACTTGGCAATGTCCTCTGAT 1362
Qy 1481 GGCACAGCTATTGTTAACTGGAGCTGGAGATGAACGCTTGTAGTTCTTGGAAATGTTT 1536
Db 1363 GGGGAGGCCATCGTCACTGGTGTGAGAGAGAGACCTGAGGTCTTGTGAAACGCTCT 1418

Search completed: January 23, 2004, 19:53:10
Job time : 7472 secs

FEATURES
source Location/Qualifiers
1..1491
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..1491
/gene="fzr1"
1..1491
/note="Fizzy-related protein"
/codon_start=1
/product="Fzr1"
/protein_id="BAA86954.1"
/db_xref="GI:6463679"
/translation="MDQYERLLRQIVIQENTMPVTEMRRLTLPASSPVSPSKH
GDRFPSPAGANWSYNFRINENKSPQNRKAKDSDNGKGLATSALEKNELIGA
GIEKVDQPTEDRLQPSPEKGLFTYLSLTKRSSPDGNDVSPYSLSPVNSKSL
LRSPPKPTKISKIPKVLDAPELDQDFYLNVLVWSSNLVSLVGLTGVYLSACTSQ
VTRLCDLSEVSDSVTSVCSWSEGNLAVGTGKHGVQIWDAAAGKLSMLHGHARVGA
LAWNAEQLSSSRDRLQIRDIRPPLQSERRLQHRQVCGKLVSTDRHLLASGGND
NKLIVWNHSSLSPVQYTHLAAVKAJAWSPHQGLLASGGTADRCIRFWNTLTGDP
LQCIDTGSQVNLAWSKHANELIVSTHYSQNLVWKYPSLTQVAKLTGHSYRVLYLA
MSPDGEAIVTGAGDETLRFNWFSTRSTKVKWSVSLNLFTRIR"
BASE COUNT 312 a 479 c 452 g 248 t
ORIGIN
Query Match 17.0%; Score 340.8; DB 9; Length 1491;
Best Local Similarity 62.8%; Pred. No. 4.2e-69;
Matches 563; Conservative 0; Mismatches 327; Indels 6; Gaps 2;
Qy 647 CTTTATAAGGTTTGGATGACCTGCTTTGCAAGATGATTTTATCTGAATCTGGTAGAT 706
Db 523 CCCTTCAAGGTGCTGGACGCGCCGAGCTGCAGGACGACTTCTACCTCAATCTGGTGGAC 582
Qy 707 TGGTCTTACACAATGTGTGCTGTGTTGGTGAATGCTGCTATCTCTGGAATGCT 766
Db 583 TGGTGTCTCCTCAATGTGCTAGCGTGGGGCTAGGCACCTCGGTGTACTGTGGAGTCCC 642
Qy 767 TGTAGCAGCAAGTAACTAAATTTATGTGA---TTTGGGGTTGATGATTTGTTGTTCT 823
Db 643 TGTACAGCAGGTGACGCGGCTCTGTGACCTCTCAGTGAAGGGGACTCAGTGACCTCC 702
Qy 824 GTTGGTTGGGCTCAAGTGGTACTCATCTGTCTGTGTAAGTAACTAACTGGTAAAGTTGAG 883
Db 703 GTGGGCTGTCTGAGCGGGGAACTTGGTGGGGTGGGCACACACAAGGGCTTCGTGGCAG 762
Qy 884 ATTTGGGATGACAGATGCAAGAGATAAGATCAATGGAGGGCCATCGTTACGTGTC 943
Db 763 ATCTGGGACGACCCGACGGAGAGAGCTGCCATGTTGGAGGGCCACACGGCACGCGTC 822
Qy 944 GGGGCTTGGCTGGAGTTCACTCTTTTGTCTTCTGGTGGACGGGATGAGATATTTAT 1003
Db 823 GGGGCGCTGGCTGGAATGCTGAGCAGCTGTCTCGGGAGCGCGACCGCATGATCCTG 882
Qy 1004 CACAGGATATACGCA---CACAGAAGATTTTGTAGTAACTGTGAGGACACAAATCA 1060
Db 883 CAGAGGACATCCGACCCCGCCACTGCGAGTGGAGCGGGCTGCAGGSCCACCAGGAC 942
Qy 1061 GAGGTTTGTGACTGAAGTGGTATGATGATAACCGTGAATGGCATCTGGAGGAAATGAC 1120
Db 943 GAGGTGTGGGGCTCAAGTGGTCCACAGACCACTCTCTCGCTCGGGGGGCAACGAC 1002
Qy 1121 AACAAATTTGTTTGGATCAACACTCAACCCAGCCTGCTCAAGTACTGTGAGCAC 1180
Db 1003 ARCAAGCTGTGTGGAATCACTGAGCCTGAGCCCTGTGAGCAGATACAGGAGCAC 1062
Qy 1181 ACAGCAGCTGTTAAAGCTATTGATGCTCTCTCATCTTTCATGGAATCTTCTGATCTGGA 1240
Db 1063 CTGGCGGCGCTGAAGGCAATCGCTGTGTCACATCAGCACGGGCTGTGGCCTCGGG 1122
Qy 1241 GAGGAACTGCAGATGATGATTTGTTTGAATACACCAACTCACACCTTAGC 1300

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 23, 2004, 10:49:52 ; Search time 596 Seconds
(without alignments)
9085.692 Million cell updates/sec

Title: US-09-701-572-1
Perfect score: 2006
Sequence: 1 gattcgacaggaagaa.....agaaaaaaaaaaaaa 2006

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2006	100.0	2006	21	AAZ37024
2	339.2	16.9	5905	21	Human ORF3013
3	334.8	16.7	2645	23	Drosophila melanog
4	303.8	15.1	1794	24	ABZ32308
5	269.6	13.4	3940	23	AA577749
6	263.8	13.2	1398	25	ABT18274
7	263.8	13.2	1398	25	ABT18868
8	263.8	13.2	1875	25	ABT20688

9	263.8	13.2	1993	25	ABT20090
10	263.8	13.2	3398	25	ABT17680
11	263.8	13.2	3393	25	ABT13494
12	259	12.9	14283	23	ABL05492
13	248	12.4	1369	23	ABL18423
14	248	12.4	3369	23	ABL18422
15	248	12.4	5058	23	ABL05448
16	231.4	11.5	1706	23	ABL07411
17	213.6	10.6	1344	24	ABZ13183
18	213.2	10.6	3852	23	ABL07410
19	200.4	10.0	1767	16	AAQ96099
20	198.8	9.9	1686	24	ABT11002
21	198.8	9.9	1686	24	ABL65393
22	198.8	9.9	1686	24	ABL65837
23	198.8	9.9	1686	24	ABL67656
24	198.8	9.9	1686	24	ABK35539
25	198.8	9.9	1688	21	AAK77801
26	198.8	9.9	1700	16	AAQ96100
27	191.4	9.5	2697	21	AAK46256
28	174.8	8.7	650	21	AAF11864
29	157.2	7.8	14482	23	ABL10086
30	152.6	7.6	3404	23	ABL06722
31	147	7.3	274	25	ABX83461
32	132.2	6.6	574	23	ABX77748
33	131.6	6.6	645	24	ABN75375
34	113.4	5.7	1034	25	ABZ82891
35	109.6	5.5	480	24	ABL81366
36	109.6	5.5	672	24	ABT09540
37	101.2	5.0	416	21	AAH30586
38	97.6	4.9	484	24	ABL82189
39	92.8	4.6	1905	24	ABZ32237
40	92.6	4.6	457	24	ABL81347
41	90.8	4.5	469	24	ABZ35113
42	78.8	3.9	678	25	ABX77642
43	70	3.5	480	24	ABL81060
44	67.8	3.4	310	16	AAZ23988
45	66.2	3.3	486	24	ABN77135

ALIGNMENTS

RESULT 1

AAZ37024
ID AAZ37024 standard; cDNA; 2006 BP.
XX
AC AAZ37024;
XX
XX 13-MAR-2000 (first entry)
XX cDNA encoding an alfa fzf protein designated ccc52Ms.

XX Alfalfa; ccc52Ms; WD40 motif; fizzy-related; fzf subfamily; fzf protein;
cellular differentiation; endoreplication; cell proliferation;
plant cell; somatic embryogenesis; endoploidy; reserve tissue;
in vitro plant regeneration; ss.

XX Medicago sativa.

XX Key Location/Qualifiers
FH Key 182..1609
FT CDS /*tag= a
FT /product= "ccc52Ms"

XX FR2779433-Al.

XX 10-DEC-1999.

XX 08-JUN-1998; 98FR-0007174.

XX 08-JUN-1998; 98FR-0007174.

XX (CNRS) CNRS CENT NAT RECH SCI.

QY 1681 TGAGCGGAGAACACCATGCTGGAAACCTTGAATATATAAAACACCAACCTAGCATC 1740
DB 1681 TGAGCGGAGAACACCATGCTGGAAACCTTGAATATATAAAACACCAACCTAGCATC 1740
QY 1741 TTTACCACTGGGAGAGCTTGGAGGAGCTATATAAGTTTGTATATGGCTGCGGTGAT 1800
DB 1741 TTTACCACTGGGAGAGCTTGGAGGAGCTATATAAGTTTGTATATGGCTGCGGTGAT 1800
QY 1801 ATTCTGCAATCATGTGTAGTCTCATTTTATTTGAAAGATGATACAAATGGTAAAT 1860
DB 1801 ATTCTGCAATCATGTGTAGTCTCATTTTATTTGAAAGATGATACAAATGGTAAAT 1860
QY 1861 TATTCCTTGGACTTATACATGATGATGAGTTGTAGCAAGTTTATTTTATTTACTCTT 1920
DB 1861 TATTCCTTGGACTTATACATGATGATGAGTTGTAGCAAGTTTATTTTATTTACTCTT 1920
QY 1921 TTTTCTTCTCTTTTGTAGTCTCTCTCTGCAATTTATTTATTTTAAATGCGT 1980
DB 1921 TTTTCTTCTCTTTTGTAGTCTCTCTCTGCAATTTATTTATTTTAAATGCGT 1980
QY 1981 TTAACAGAAAAAATAAAAAA 2006
DB 1981 TTAACAGAAAAAATAAAAAA 2006

RESULT 2

AACT7458
ID AACT7458 standard; cDNA; 5905 BP.
XX AACT7458;
AC AACT7458;
DT 08-FEB-2001 (first entry)
DE Human ORF3013 polynucleotide sequence SEQ ID NO: 6025.
XX Human; open reading frame; ORF; detection; cytostatic; hepatotropic;
KW vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antineoplastic;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; diabetes mellitus; graft vs host disease;
KW cardiovascular disease; systemic lupus erythematosus; infection;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive; ss.

XX Homo sapiens.
OS
XX WO200058473-A2.
XX
XX '05-OCT-2000.
XX
XX 31-MAR-2000; 2000WO-US08621.
XX
XX 31-MAR-1999; 99US-0127607.
XX 02-APR-1999; 99US-0127636.
XX 05-APR-1999; 99US-0127728.
XX 30-MAR-2000; 2000US-0540763.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Shimketa RA, Leach M;
XX
XX WPI; 2000-602362/57.
XX P-PSDB; AAB43249.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
XX useful for treating e.g. cancers, proliferative disorders,
XX neurodegenerative disorders and cardiovascular disease -

XX Claim 5; Page 5206-5209; 5507pp; English.
PS
XX AACT74446 to AACT7606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORF open reading frames 1 to 3161. The ORF
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antithyroid; and antianemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORF-associated disorder. The
CC nucleic acids can be used to express ORF proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, diabetes mellitus,
CC graft vs host disease, cardiovascular disease, osteoarthritis,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ Sequence 5905 BP; 1095 A; 1770 C; 1963 G; 1076 T; 1 other;

Query Match 16.9%; Score 339.2; DB 21; Length 5905;
Best Local Similarity 62.7%; Pred. No. 1.8e-68;
Matches 562; Conservative 0; Mismatches 328; Indels 6; Gaps 2;
QY 647 CCTATATAGTTTGGATGCACCTGTTTGAAGATGATTTTATCTGAATCTGGTAGAT 706
DB 784 CCCTTAGGTGCTGAGCGGCCGAGCTGCAGGACGACTTCTACCTCAATCTGGTGAC 843
QY 707 TGGTCTTACAAATGTTGGCTGTGGTTGGTAACTGTGTCTATCTCTGGAATGCT 766
DB 844 TGGTGTGCTTCAATGTCTCAGCTGGGGCTAGGCACTGCTGTACCTGTGGAGTGC 903
QY 767 TGTAGCAGCAAGTAACTAAATATATGA---TTTGGGGCTTGTATGTTGTTTCT 823
DB 904 TGTACCAAGCAGGTGACCGGGCTCTGTGACCTCTCAGTGAAGGAGACTCAGTGACCTCC 963
QY 824 GTTGGTTGGGCTCAACGCTGTACTCATCTTGTGTTTGGAACTAACTAAATGTAAGTTTCAG 883
DB 964 GTGGGCTGGTCTGAGCGGGGAACTGCTGGCGTGGGACACACAAAGGGCTTCGTGCAG 1023
QY 884 ATTTGGGATCGCAGCAGATCGAAGATCAATGAGGGCCATCGTTCAGCTGTC 943
DB 1024 ATCTGGGACGCGCCGAGGAAAGAGCTGTCCAATGTTGGAGGGCCACACGACCGCTC 1083
QY 944 GGGGCTTGGCTGGAGTTTCACTCTTTTCTTCTTGTGCGAGGGGATAGAATATTTAT 1003
DB 1084 GGGGCTTGGCTGGAGTTTCACTCTTGTGCGAGCTGTCTGCGGAGCGCGACCATGCTG 1143
QY 1004 CAACGAGATATACGCA---CACAAAGAGATTTTGTAGTAACTGTGAGGACACAAATCA 1060
DB 1144 CAGAGGACATCCGACCCCGCCACTGCTGCTGAGCGGGCTGTCAGGGGCCACCGGAC 1203
QY 1061 GAGTTTGTGAGCTGAAGTGGTCAATATGATAACCGTGTGGTGGCATCTGGAGGAAATGAC 1120
DB 1204 GAGTTTGTGAGCTGAAGTGGTCAATATGATAACCGTGTGGTGGCATCTGGAGGAAATGAC 1263
QY 1121 AACAAATTTTGTGGAATCAACACTCAACCCAGCGCTGTCTCTCAATGTTGAGTCTGTGAGCAC 1180
DB 1264 AACAAATTTTGTGGAATCACTCGAGGCTGAGCGGCTGTGAGGAGTACACGAGCAC 1323
QY 1181 ACAGCAGCTGTTAAAGCTTATGATGCTCTCTCATCTTGTGAGTCTTGTGATCTTGGGA 1240
DB 1324 CTGGCGGCGGTGAAGGCTATGCTGCTCTCCCATCATCAGCAGGGCTGTGGCTCGGG 1383
QY 1241 GGAGAACTGCAGATAGATGATTTGCTTTTGGAAATCAACCAAACTCACACTTAC 1300

Db 1384 GGGGACAGCTGACCGCTGTATCCGCTTCTGGACACAGCTGACAGGACCAACCACTGCG 1443
 QY 1301 TGTATGGACACTGGAAAGTCAGTTTGCAATCTTCTGTGCTCCAAAATGTCACAGCACTA 1360
 Db 1444 TGTATCGACACGGGCTCCCAAGTGTGCAATCTGCGCTGGTCCAAAGCAGCGCAAGAGCTG 1503
 QY 1361 GTAAGCACACATGGGTACTCCAGAACACAGATTATTGTTTGGAGATACCCCACTATGTCA 1420
 Db 1504 GTGAGCAGCAGCGGTACTCTACAGAACAGATCTCTGTCTGGAAGTACCCCTCCCTGACC 1563
 QY 1421 AGCTGGGACTCTTACCGGCCATATTATAGGTTCTCTATCTTGCCATCTCTCCAGAT 1480
 Db 1564 CAGGTGGCCAACTGACCGGGCACTCTACCGCGTGTGTACCTGGCAATGTCCCTGTAT 1623
 QY 1481 GGACAGACTATTGTAACTGGAGCTGGAGATGAAACGCTTATAGTTCTCGAAATGTTTT 1536
 Db 1624 GGGGAGGCATCTGCTCACTGGTGTGAGACAGAGACCCCTGAGGTTCTCGAAAGCTTT 1679

RESULT 3

ABL05493
 ID ABL05493 standard; cDNA; 2645 BP.

XX ABL05493;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 10961.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical; gene; ss.

XX Drosophila melanogaster.

PN WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEXE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX P-P8DB; ABB61390.

XX New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -

XX Claim 1; SEQ ID NO 10961; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072).

XX The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 2645 BP; 766 A; 630 C; 619 G; 630 T; 0 other;

XX Query Match

XX Best Local Similarity 16.7%; Score 334.8; DB 23; Length 2645; Mismatches 572; Conservative 0; Mismatches 352; Indels 6; Gaps 2;

QY 614 CCGTTAAGGCTCCTAGGAAGGTTCTCTGATCCGCTTATAAGGTTTGGATGCACTGCT 673
 Db 942 CCGCGCAAGGCTACGCGCAAAATCTCTCGCAATTCCTTCAAGGTGTAGACGCGCCGAG 1001
 QY 674 TTGCAAGATGATTTTATCTGAATCTGGTAGATTGGTCTTTCACCAATGTTGGTGT 733
 Db 1002 TTGCAAGACGACTTCTATCTGAACCTGGTCTGCTGCTCGCAGAACGTTACTGGCTGA 1061
 QY 734 GTTTGGGTAACCTGCTCTATCTCTGGAATGCTGTAGCAGCAAGTAATAATTATCT 793
 Db 1062 GCGCTGGGCACTGCTGCTATCTGTGGAGCGCGTGCACCACTCAGTTTACCGCCCTGTGT 1121
 QY 794 GATTGGG---GGTTGATGATTGTGTTTCTGTTGGTCTCAAGTGTGCTACTCAT 850
 Db 1122 GATCTCAGTCCGATCGAATACGTTGACCTCGTGTCTGTGGAAACGAGGTGGCAACAC 1181
 QY 851 CTTGCTGTTGGAACCTAATGTTAAAGTTACATTTGGATGCGACAGATGCAAGAG 910
 Db 1182 GTGCGCGTGGGCAACATCACCGCTAGTGACCGTCTGGGATGTGCGGCCAATAAGCAG 1241
 QY 911 ATAAGATCAATGGAGGCGCATCGTTTACGTGTGCGGGCCCTTGGCCCTGGAGTTTCACTCT 970
 Db 1242 ATCAACAACTGAATGGCCATTCGCGCGTGTGGCGCCCTTGGCATGGAACAGTGCATC 1301
 QY 971 TTGCTTCTGCTGAGCGGATAGAATATTATCAACGAGATATACGCA---CACAAGA 1027
 Db 1302 CTGTGAGCGGTGCGGAGACCGTTGGATCATACACGGGATACGAGAACGCGCACTG 1361
 QY 1028 GATTTGTAGTAACTGTGAGACACAAATCAGAGGTTTGTGGACTGAAAGTGTGTCATAT 1087
 Db 1362 CAATCGGAGCGCATGATTGGCGGACATCGCGAGAGGTGTGCGGACTGAAATGTTCA 1421
 QY 1088 GATAACCGTGGATTGGCATCTGGAGGAATGACAAATTTGTTTCTTGGGATCAACAC 1147
 Db 1422 GATTAATCAATCTTGGCCAGTGGCGCAACGATATCGGTTGATGTGGATCAGCAT 1481
 QY 1148 TCAACCCAGCGCTGCTCAAGTACTGTGAGCACACAGCAGCTGTTTAAAGCTATTGCA 1207
 Db 1482 TCCGTGAATCCGCTACAAATCATACAGGAGCATATGCGCGCTGTAAAGGCGATCGG 1541
 QY 1208 TCTCTCATCTTCATGACTTCTTGGCATCTGGAGGAGGAAGTGCAGATAGATGTTTCT 1267
 Db 1542 TCGCGCATCACACAGGACTCTTGGCCAGCGCGGTGGAAACGCGGATAGGTGTTATCC 1601
 QY 1268 TTTTGAATACCAACCAAACTCACACCTTAGCTGTATGGACACTGGAAGTCAGGTTTC 1327
 Db 1602 TTCTGGAATACGCTGACCGGCGCAGCCCATGCACTGGGTGGACACGCGGCTCG 1661
 QY 1328 AATCTGTGCTGGTCCAAAATGTCAAGAACTAGTAGCACACATCGGTTACTCCCAAG 1387
 Db 1662 AATCTGCGCTGGTCCAAGCACTCTCTCGGAGCTGCTCTCCAGCACGCGCTACT 1721
 QY 1388 CAGATTATTGTTGGAGATACCCCACTATGTCAAGCTGGGACTCTTACCAGCCATAT 1447
 Db 1722 CAGATCTCTGTGGAAATATCCCTCCCTGACCAAGTGGCCCAAGCTGACGGGCCA 1781
 QY 1448 TATAGGTTCTCTATCTTGGCCATCTCTCCAGATGGACAGACTATTCTTAAGTGGAG 1507
 Db 1782 TATGCTGTGCTCTATCTGGCGGTGAGTCCGATGGTGGAGGCTATTGTTTACGG 1841
 QY 1508 GATCAAAACGCTTAGTGTCTCGAATGTTTC 1537
 Db 1842 GACGAGACGCTGCGAATTTTGAACGTATTC 1871

RESULT 4

ABZ32308
 ID ABZ32308 standard; DNA; 1794 BP.

XX AC ABZ32308;

XX DT 30-JAN-2003 (first entry)

XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Dmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
XX DR P-PSDB; ABG13562.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity -
XX PS Claim 1; SEQ ID No 13553; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC (II). (II) is useful for generating antibodies against it, detecting or
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. AAS64197-AAS94564 represent novel human
XX CC diagnostic coding sequences of the invention.
XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 3940 BP; 799 A; 1246 C; 1221 G; 674 T; 0 other;
Query Match 13.4%; Score 269.6; DB 23; Length 3940;
Best Local Similarity 62.6%; Pred. No. 2.1e-52;
Matches 438; Conservative 0; Mismatches 259; Indels 3; Gaps 1;
QY 840 GTGGTACTCATCTTCTGCTGGAATCAATGTAAGTTCAGATTGGATGACGAA 899
DB 2357 GGGGGAACCTGGTGGCGGTGGGACACACAAAGGGCTTCGTGCAGATCTGGAGCAGCG 2416
QY 900 GATGCAAGATAGATCAATGAGGGCCATCGTTACGTGTCGGGCTTGGCCCTGGA 959
DB 2417 CAGGGAAGAAGTGTCCATGTTGGAGGGCCACACGGCAGCGTTCGGGCGCTGGCTTGA 2476
QY 960 GTTCATCTCTTTTCTTCTGTCGACGGGATAGAAATATTATCAACAGATATACGA 1019
DB 2477 ATGCTGAGCAGCTGCTGCTGGGAGCGCGGACCGCATCATCTCTGACAGAGGACATCGCA 2536
QY 1020 ---CAACAGAAGATTTTGTAGTAACTGTCAGGACAAATCAGAGTTTGTGACTGA 1076
DB 2537 CCCCCCACTGAGTCGAGCGGCGGCTGACGGGCGCACCGGAGAGGTGTGGCGGTCA 2596
QY 1077 AGTGGTCATATCAATCACTGAGTTGGCATCTGGAGGAATGACAAATATTTTGT 1136
DB 2597 ATGTTGTCACAGACCACTGCTCTGCTCGCTCGGGGGCAACGACAAAGCTGTGGTCT 2656
QY 1137 GGAATCAACATCAACCCAGCTGCTCTCACTGTCAGGACACACAGCTGTAAAG 1196
DB 2657 GGAATCACTCAGAGCTGAGCCCGGTGCAGCAGTACACGGAGCACCTGGCGCGCGTGAAGG 2716

QY 1197 CTATTGATGGTCTCTCATCTTCTGACATCTTCTGACATCTGAGGAGGAACTGCAGATA 1256
DB 2717 CCATCGCTGCTCCCAATCAGCACGGGCTGTGGCTCGGGGGGGGGGCGGACAGCTGACC 2776
QY 1257 GATGATTTGCTTTTGGAAATACAAACCAAACTCAGACCTTAGCTGTATGGACACTGGAA 1316
DB 2777 GCTGTATCCGCTTCTGGAACACGCTGACAGGACAACTGCACTGTATGACACGCGCT 2836
QY 1317 GTCAGGTTGCAATCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1376
DB 2837 CCCAAGTGTGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2896
QY 1377 ACTCCAGAACCAAGATATTGTTTGGAGATACCCCACTATGTCAAAGCTGGCGACTCTTA 1436
DB 2897 ACTACAGAACCAAGATCTTGTCTGGAAGTACCCCTCCCTGACCCAGCTGGCCAACTGA 2956
QY 1437 CCGGCAATCTTATGAGGTTCTTATCTTGGCATCTCTCCAGATGGACAGACTATTGTAA 1496
DB 2957 CCGGCACTCTTACCGCTGCTGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3016
QY 1497 CTGGAGCTGGAGATGAACAGCTTAGGTTCTGGAATGTTT 1536
DB 3017 CTGGTCTGGAGACGAGACCCCTGAGTTCTGGAAGCTCTT 3056
RESULT 6
ABT18274
ID ABT18274 standard; DNA; 1398 BP.
XX AC ABT18274;
XX DT 16-APR-2003 (first entry)
XX DE Aspergillus fumigatus essential gene #632.
XX KW Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;
XX KW cancer; contamination; biofilm; antibody; immune response; ds.
XX OS Aspergillus fumigatus.
XX PN WO200286090-A2.
XX PD 31-OCT-2002.
XX PF 23-APR-2002; 2002WO-US13142.
XX PR 23-APR-2001; 2001US-285697P.
XX PR 27-APR-2001; 2001US-287066P.
XX PR 05-JUN-2001; 2001US-295890P.
XX PR 09-JUL-2001; 2001US-303899P.
XX PR 31-AUG-2001; 2001US-316362P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Jiang B, Tishkoff D, Zamudio C, Broshkin AM, Hu W, Lemieux SM;
XX DR WPI; 2003-093124/08.
XX PT New purified or isolated nucleic acids of essential genes of
XX PT Aspergillus fumigatus, useful for treating or preventing infections by
XX PT A. fumigatus, or for treating a non-infectious disease in a subject
XX PT e.g. cancer -
XX PS Disclosure; Page -; 175pp; English.
XX CC The invention relates to novel purified or isolated nucleic acids of
XX CC essential genes of Aspergillus fumigatus. The isolated nucleic acids of
XX CC the invention are used to treat or prevent infections by a pathogenic
XX CC organism such as A. fumigatus, to treat a non-infectious disease in a
XX CC subject (e.g. cancer), to prevent or contain contamination of an object
XX CC by A. fumigatus, or to prevent or inhibit formation on a surface of a
XX CC biofilm comprising A. fumigatus. The polynucleotides are useful for

CC expressing recombinant protein for characterisation, screening or
 CC therapeutic use, as markers for host tissues in which the pathogenic
 CC organisms invade or reside, for comparing with the DNA sequence of A.
 CC fumigatus to identify duplicated genes or paralogues having the same or
 CC similar biochemical activity and/or function, for comparing with DNA
 CC sequences of other related or distant pathogenic organisms to identify
 CC potential orthologous essential or virulence genes, for selecting and
 CC making oligomers for attachment to a nucleic acid array for examination
 CC of expression patterns, for raising anti-protein antibodies, as an
 CC antigen to raise anti-DNA antibodies or to elicit another immune
 CC response, and for identifying polynucleotides encoding the other protein
 CC with which binding occurs or to identify inhibitors of the binding
 CC interaction. The polypeptides may be used to raise antibodies or to
 CC elicit immune response, as a reagent in assays designed to quantitatively
 CC determine levels of the protein in biological fluids, as a marker for
 CC host tissues in which pathogenic organism invade or reside, and to
 CC isolate correlative receptors or ligands in the case of virulence
 CC factors. This polynucleotide sequence represents one of the essential
 CC genes of *Aspergillus fumigatus* of the invention.
 XX
 SQ Sequence 1398 BP; 368 A; 360 C; 343 G; 327 T; 0 other;

Query Match 13.2%; Score 263.8; DB 25; Length 1398;
 Best Local Similarity 56.3%; Pred. No. 3.4e-51;
 Matches 557; Conservative 0; Mismatches 372; Indels 60; Gaps 1;

QY 609 ATAGTCCGGTTAAGGCTCCTAGAAAGGTTCTCGATCCCTTATAAGGTTTGGATGCAC 668
 DB 341 AGACTCTCGCAAGCAGCGCTCGTACGTTAATAAGTACCTTATAAGGTTCTCGACGCAC 400

QY 669 CTGCTTTCAGATGATTTTATCTGAATCTGGTAGTGTCTTCACACATGTTGG 728
 DB 401 CGGATTTGCAAGATGATTTCTACTTGAATCTGGTAGTGTGGCAGTAGTAAATGTTAG 460

QY 729 CTGTTGGTTTGGGTAACTGTGTCTATCTCTGGAATGCTTGTAGCAGCAAGTAATAAT 788
 DB 461 GGTGTGGCTAGCAATTCAGTGTACATGTGGAAATTCGAATACCGACGGGTGACGAAC 520

QY 789 TATGTGATTTGGGGTTGATGTTG-----TGTTTGTCTGTTGGTCTCAAGTGTGATCTC 813
 DB 521 TTTCGCACTAAGAGATGATGATGTCACAAAGTTAGTGTGATTCAGAGGGTATAGAT 580

QY 814 -----TGTTTGTCTGTTGGTCTCAAGTGTGATCTC 848
 DB 581 TTCTCTATATGTTGGCAATGATTAATATGTTTCTAGCTAGCATTCGACAGGGCACAC 640

QY 849 ATCTGTCTTGGTGAACCTAACAATGTTAAAGTTTCAATTTGGATTCAGCAGCAAGTCAAGA 908
 DB 641 ATCTTTCAATAGGAACCTGCAAAAGGTTTGTACAGATATGGGATCGAGCAATTTGCGCC 700

QY 909 AGATAAGATCAATGGAGGCGCCATCGTTACGTGTGGGGCTTGGCCCTGGAGTTTCATCTC 968
 DB 701 GTCTTCGGACATGATTTGGGCAATACCAATCGTGTGGGGCTCTTCTTGGACGATCATA 760

QY 969 TTTTGTCTTCTGGTGAACGGGATAGAAATATTATTAACAGAGATATACGACACAAGAG 1028
 DB 761 TCTTACGTCAGGTTCTCGGATCGACTGATCTCCACCGTATGTTGTTTCCCGAGATC 820

QY 1029 ATTTGTTAGTAACTGTGACGACCAATCAGAGTTTGTGAGTGAAGTGTCTATATG 1088
 DB 821 AGTATTACGTGACTGTCCGGCCATAACGAGAAAGTTTCGGATTCGGTGGAAACCG 880

QY 1089 ATAACCGTGAATGGCATCTGAGGAAATGACAAACAATTTGTTTGGAAATCAACT 1148
 DB 881 AAGATGGCCAACTAGCTTCAGGCGGAATGACAAACAATCTGATGTTGGGACAGTGA 940

QY 1149 CAAACCGGCTGCTCAAGTACTGTGACACACAGCAGCTGTTAAAGCTTATGATGTT 1208
 DB 941 ATGAGACACCGTTTATTCGCTTCTCGACCATACCGACCGGTGAAGGCCATCGGCTGTT 1000

QY 1209 CTCTCATCTTCAATGACTTCTTGTGACTCTGGAGGAGAACTGCAGATAGATGATTTGTT 1268
 DB 1001 CACCCCATCAACACCATCTCTCGCTCAGTGGAGGACGCGGACCGACCATCAAT 1060

RESULT 7

ABT18868

ID ABT18868 standard; DNA; 1398 BP.

XX AC ABT18868;

XX XX 16-APR-2003 (first entry)

XX DE *Aspergillus fumigatus* essential gene #1226.XX XX Fungicide; cytostatic; essential gene; *Aspergillus fumigatus*; infection;
 XX KW cancer; contamination; biofilm; antibody; immune response; ds.XX OS *Aspergillus fumigatus*.

XX XX WO200286090-A2.

XX PD 31-OCT-2002.

XX XX 23-APR-2002; 2002WO-US13142.

XX XX 23-APR-2001; 2001US-285697P.

XX PR 27-APR-2001; 2001US-287066P.

XX PR 05-JUN-2001; 2001US-295890P.

XX PR 09-JUL-2001; 2001US-303899P.

XX PR 31-AUG-2001; 2001US-316362P.

XX PA (ELIT-) ELITRA PHARM INC.

XX XX Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;

XX XX WPI; 2003-093124/08.

XX XX New purified or isolated nucleic acids of essential genes of
 XX XX *Aspergillus fumigatus*, useful for treating or preventing infections by
 XX XX A. fumigatus, or for treating a non-infectious disease in a subject
 XX XX e.g. cancer

XX PS Disclosure; Page -: 175pp; English.

XX CC The invention relates to novel purified or isolated nucleic acids of
 XX CC essential genes of *Aspergillus fumigatus*. The isolated nucleic acids of
 XX CC the invention are used to treat or prevent infections by a pathogenic
 XX CC organism such as A. fumigatus, to treat a non-infectious disease in a
 XX CC subject (e.g. cancer), to prevent or contain contamination of an object
 XX CC by A. fumigatus, or to prevent or inhibit formation on a surface of a
 XX CC biofilm comprising A. fumigatus. The polynucleotides are useful for
 XX CC expressing recombinant protein for characterisation, screening or
 XX CC therapeutic use, as markers for host tissues in which the pathogenic
 XX CC organisms invade or reside, for comparing with the DNA sequence of A.
 XX CC fumigatus to identify duplicated genes or paralogues having the same or

CC similar biochemical activity and/or function, for comparing with DNA
CC sequences of other related or distant pathogenic organisms to identify
CC potential orthologous essential or virulence genes, for selecting and
CC making oligomers for attachment to a nucleic acid array for examination
CC of expression patterns, for raising anti-protein antibodies, as an
CC antigen to raise anti-DNA antibodies or to elicit another immune
CC response, and for identifying polynucleotides encoding the other protein
CC with which binding occurs or to identify inhibitors of the binding
CC interaction. The polypeptides may be used to raise antibodies or to
CC elicit immune response, as a reagent in assays designed to quantitatively
CC determine levels of the protein in biological fluids, as a marker for
CC host tissues in which pathogenic organism invade or reside, and to
CC isolate correlative receptors or ligands in the case of virulence
CC factors. This polynucleotide sequence represents one of the essential
CC genes of *Aspergillus fumigatus* of the invention.
XX
SQ Sequence 1398 BP; 368 A; 360 C; 343 G; 327 T; 0 other;

Query Match 13.2%; Score 263.8; DB 25; Length 1398;
Best Local Similarity 56.3%; Pred. No. 3.4e-51;
Matches 557; Conservative 0; Mismatches 372; Indels 60; Gaps 1;

QY 609 ATAGTCGGTTAAGGCTCCTAGAGAGGTTCCGATCGCCCTTATAAGGTTTGGATGCAC 668
DB 341 AGACTCCTCGACAGCCCTCGTACGTTAATAAAGTACCCTATAAGGTTCTCGACGCAC 400

QY 669 CTCCTTTGCAAGATGATTTTATCTCAATCTGATAGTTGGTCTTCACACATGTGTGG 728
DB 401 CCGATTTGCAGATGATTTCTACTTGAATCTGGTGCATGGGCGAGTAGTATGTCTAG 460

QY 729 CTGTTGGTTGGTAACTGTGTCTATCTCTGGAATCTGTAGCAGCAAGGTAACTAAT 788
DB 461 GTGTTGGCTTAGGCAATTCAGTTACATGTGGAATTCGAATACCGGACGGGTGACGAAC 520

QY 789 TATGTGATTTGGGGTTGATGTTG----- 813
DB 521 TTTGCGAATTAAGATGATGATCTGTCAGAGTTAGTGGATTACAGGGTAAATGAGAT 580

QY 814 -----TGTTTGTCTGTTGGTTGGCTCAACGTGGTACTC 848
DB 581 TTCTCTATATGTTGGCAATGATTTATGTTTCTATGATCTGACATTTGGACAGGGCACAC 640

QY 849 ATCTTGCTGTGAGCAATGAAGTAAAGTTGAGATTTGGGATGACGCAAGATGCAAGA 908
DB 641 ATCTTTCAATAGGAATCGGCAAGGCTTGTACAGATATGGGATGACAGCATTTGCGCC 700

QY 909 AGATAAGATCAATGAGGGCCATCGGTTACGTCTCGGGGCTTGGCTGGAGTTCACTC 968
DB 701 GTCTTCGACAAATGATTGGGCATACCAATCGTGTGGGGCTCTTGTGGAAAGATCATA 760

QY 969 TTTTGTCTTCTGGTGGACGGGATAAGATATTTATCAACGAGATATACGACACAGAAAG 1028
DB 761 TCCTTACGTAGGTTCTCGGGATCGACTGATCTTCCACCGTGATGTTGTTCCCGCAGATC 820

QY 1029 ATTTTGTAGTAACTCTCAGGACACAAATCAGAGGTTTGTGGACTGAAGTGTCTATATG 1088
DB 821 AGTATTACGTGCACTCTCGGCCATTAAGCAGGAAGTTTCGGGACTCGGTGAACACCG 880

QY 1089 ATAACCGTGAATTGGCACTGGAGGAATGACAAATTTGTTTGGTGAATCAACACT 1148
DB 881 AAGATGCGCAACTAGCTTTCAGGCGGAAATGACAACTCATGGTGTGGGACAGCTGA 940

QY 1149 CAACCCAGCCTGCTCCTCAAGTACTGTGACGACACAGCAGCTGTTAAAGCTATTGCTATGGT 1208
DB 941 ATGAGACACCGCTTATTCGCTTCTCGACCATACCGGACCGGTGAAGCCATCGCTGGT 1000

QY 1209 CTCCTCATTTTATGAGATCTTTCGATCTGTGAGGAGGAATCGAGATAGATGATTCGTT 1268
DB 1001 CACCCCATCAACACCATCTCTCCCTCAGTGGAGGACGGGCGACCGGACCATCAAT 1060

QY 1269 TTTGGAATACACCAACCACTCAACCTTAGCTGATTTGGACACTGGAAGTCAAGTTGCA 1328
DB 1061 TCTGGAACATGCTACTGGTTCACTGATCAAGGAAGTTGATACGGGTAGCCAGGTGTGA 1120

QY 1329 ATCTTGTCTGGTCCAAATGTCAAGAACTAGTAGACACACATGGGTACTCCCAAGACC 1388
DB 1121 ACCTGCGATGGTCCGAAGAACTCGAGCAATCATCTAGTACACCGCTACAGTCAGAAC 1180

QY 1389 AGATTTATTTTGGAGATACCCCACTATCTCAAAAGCTGGCGACTCTTACCGGCCATACTT 1448
DB 1181 AGATTGTAACTCGAAATACCCGCGCATGGAGCAGATTGTATCGCTGACCGCGCACACTT 1240

QY 1449 ATAGGTTCTCTATCTTGGCATCTCTCCAGATGGACAGACTATTGTAACTGGAGCTGGAG 1508
DB 1241 TCCGTGTTCTCTACCTTGGCATGAGCCCGATGAGCCGATGAGCAGACCGCTCGTTCGGAGCTGGTG 1300

QY 1509 ATGAAACGCTTAGGTTCTGGAATGTTTC 1537
DB 1301 ACGAACGTTGAGATTCTGGAAGATTTTC 1329

RESULT 8
ABT20688
ID ABT20688 standard; DNA; 1875 BP.
XX
AC ABT20688;
DT 16-APR-2003 (first entry)
XX
DE *Aspergillus fumigatus* essential gene #3046.
XX
KW Fungicide; cytostatic; essential gene; *Aspergillus fumigatus*; infection;
XX cancer; contamination; biofilm; antibody; immune response; ds.
XX
OS *Aspergillus fumigatus*.
XX
FN WO200286090-A2.
XX
PD 31-OCT-2002.
XX
PF 23-APR-2002; 2002WO-US13142.
XX
PR 23-APR-2001; 2001US-285697P.
PR 27-APR-2001; 2001US-287066P.
PR 05-JUN-2001; 2001US-295890P.
PR 09-JUL-2001; 2001US-303899P.
PR 31-AUG-2001; 2001US-316362P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Jiang B, Tishkoff D, Zamudio C, Broshkin AM, Hu W, Lemieux SM;
XX WPI; 2003-093124/08.
XX
DR New purified or isolated nucleic acids of essential genes of
PT *Aspergillus fumigatus*, useful for treating or preventing infections by
PT *A. fumigatus*, or for treating a non-infectious disease in a subject
PT e.g. cancer -
XX
XX Disclosure; Page -: 175pp; English.
XX
CC The invention relates to novel purified or isolated nucleic acids of
CC essential genes of *Aspergillus fumigatus*. The isolated nucleic acids of
CC the invention are used to treat or prevent infections by a pathogenic
CC organism such as *A. fumigatus*, to treat a non-infectious disease in a
CC subject (e.g. cancer), to prevent or contain contamination of an object
CC by *A. fumigatus*, or to prevent or inhibit formation on a surface of a
CC biofilm comprising *A. fumigatus*. The polynucleotides are useful for
CC expressing recombinant protein for characterization, screening or
CC therapeutic use, as markers for host tissues in which the pathogenic
CC organisms invade or reside, for comparing with the DNA sequence of *A.*
CC *fumigatus* to identify duplicated genes or paralogues having the same or
CC similar biochemical activity and/or function, for comparing with DNA
CC sequences of other related or distant pathogenic organisms to identify
CC potential orthologous essential or virulence genes, for selecting and
CC making oligomers for attachment to a nucleic acid array for examination

CC of expression patterns, for raising anti-protein antibodies, as an
CC antigen to raise anti-DNA antibodies or to elicit another immune
CC response, and for identifying polynucleotides encoding the other protein
CC with which binding occurs or to identify inhibitors of the binding
CC interaction. The polypeptides may be used to raise antibodies or to
CC elicit immune response, as a reagent in assays designed to quantitatively
CC determine levels of the protein in biological fluids, as a marker for
CC host tissues in which pathogenic organism invade or reside, and to
CC isolate correlative receptors or ligands in the case of virulence
CC factors. This polynucleotide sequence represents one of the essential
CC genes of *Aspergillus fumigatus* of the invention.
XX

SQ Sequence 1875 BP; 514 A; 490 C; 464 G; 407 T; 0 other;

Query Match 13.2%; Score 263.8; DB 25; Length 1875;

Best Local Similarity 56.3%; Pred. No. 3.7e-51;

Matches 557; Conservative 0; Mismatches 372; Indels 60; Gaps 1;

QY 609 ATAGTCGGTTAAGGCTCCAGGAGTTCTTCGATCGCTTAAAGTTTGGATGCAC 668
Db 818 AGACTCTCGAAGCAGCTCGCTAGCTTAAATAAGTACCCCTATAAGGTTCTCGACGCAC 877
QY 669 CTGCTTTGCAAGATGATTTTATCTGAATCTGCTAGATGCTTCACACATGCTGG 728
Db 878 CCGATTTCGAAGATGATTTTCTACTTGAATCTGCTAGCTGCGGCAGTAGTAACTGCTAG 937
QY 729 CTGTTGGTTGGGTAACTGCTGCTATCTCTGGAATGCTTTAGCAGCAAGGTAACATAAT 788
Db 938 GTGTTGGCTAGGCAATTCAGTGTACATGTGGAATTCGAATACCGGCGGGTGACGAAC 997
QY 789 TAATGATTTGGGGTGTGATGTTG----- 813
Db 998 TTTCGCAACTAAGAGATGATCTGTTCACAAAGTTAGTGTGGAATCAGAGGGTAATGAGAT 1057
QY 814 -----TGTTTGTCTTGTGTTGGCTCAACGTGTGCTC 848
Db 1058 TTCTCTATATGTTGGCAATGATTAATGTTTATGTTTATGCTGACATGTCGACAGGCGACAC 1117
QY 849 ATCTGCTGTGGAATCAATGTAAGTTTGAATTTGGATGTCAGCAAGATGCAAGA 908
Db 1118 ATCTTTCAATAGGAACCTGGCAAGGCTTTGTACAGATATGGGATGCGAGCAATTCGCCC 1177
QY 909 AGATAAGATCAATGGAGGCGCATCGTTACGTTGTCGGGCGCTTGGCTGAGTTTATCTC 968
Db 1178 GTCTTCGACATATGTTGGCATACCAATCGTGTGGGGCTCTTGTGGAACGATCATA 1237
QY 969 TTTTGTCTTCTGTGTCAGGATAGAAATATTTATCAACGAGATATACGCAACAAGAG 1028
Db 1238 TCCTTACGTCAGGTTCTCGGATCGACTGATCTTCCACGCTGATGTTCTCCCGAGATC 1297
QY 1029 ATTTTGTAGTAACTGTGAGGACAAATCAGAGCTTTGTGACTGAACTGATGTCATATG 1088
Db 1298 AGTATTTACGTCAGCTGTCGGCCATAGCAGGAAGTTTGGCACTCCGGTGGACACCG 1357
QY 1089 ATAACGCTGATTTGGCATCTGGAGGAATGACAAACAAATTTGTTTGGATCAACACT 1148
Db 1358 AAGATGCCCACTAGCTTTCAGGCGGAATGACAAACAACTCATGTTGTGGGCAAGCTGA 1417
QY 1149 CAACCCAGCTGCTCCTCAAGTATCTGAGCAGACAGCAGCTGTTAAGTATGATGCT 1208
Db 1418 ATGAGACACCGCTTTATCGCTTCTCGACCATACCGACCGCTGAAAGGCGCATCGCGTGT 1477
QY 1209 CTCTCATCTTCATGCACTTCTTGCACTCTGGAGGAACTGACAGATAGATGATTCGTT 1268
Db 1478 CACCCCATCAACACCATCTCTCGCTCAGTGTGAGGCGAGCGGCGGACCATCAAT 1537
QY 1269 TTTGGAATCAACACCAAACTCACTTACCTTGTATGAGACATGGAAGTCAAGTTTGA 1328
Db 1538 TCTGGAACACTGTTACTGTTTCACTGATCAAGGAAGTTGATACGGGTAGCCAGGTGTGA 1597
QY 1329 ATCTTGTCTGTCCTCAAAAGTCAACCAACTAGTAAGCAGACATGTTGTTCTCCAGAAC 1388
Db 1598 ACCTGGCATGTTGGAAGAACTCGGACGAATCATCAGTACACACGCTACATGACAGAAC 1657

QY 1389 AGATTATGTTTGGAGATACCCACACTATGTCAAAGCTGGCGACTCTTACGGCCACTT 1448
Db 1658 AGATTGTAATCTGGAATATCCCGCATGGACGAGTATGCTGACCGGCGACACTT 1717
QY 1449 ATAGGCTTCTTCTATCTTGGCATCTCTCCAGATGACAGACTATGTAATCTGGAGCTGGAG 1508
Db 1718 TCCGTGTTCTTCTACTACCTTGGATGAGCCCGATGACAGACCGCTGTTACCGGAGCTGGTG 1777
QY 1509 ATGAAAGCTTAGGTTCTGGAATGTTTC 1537
Db 1778 ACGAACGTTGAGATTCTGGAAGATTTC 1806

RESULT 9

ABT20090

ID ABT20090 standard; DNA; 1993 BP.

XX

AC ABT20090;

XX

DT 16-APR-2003 (first entry)

XX

DE *Aspergillus fumigatus* essential gene #2448.

XX

XX Fungicide; cytostatic; essential gene; *Aspergillus fumigatus*; infection;

KW cancer; contamination; biofilm; antibody; immune response; ds.

XX

OS *Aspergillus fumigatus*.

XX

PN WO200286090-A2.

XX

PD 31-OCT-2002.

XX

PF 23-APR-2002; 2002WO-US13142.

XX

PR 23-APR-2001; 2001US-285697P.

PR 27-APR-2001; 2001US-287066P.

PR 05-JUN-2001; 2001US-295890P.

PR 09-JUL-2001; 2001US-302899P.

PR 31-AUG-2001; 2001US-316362P.

XX

PA (ELIT-) ELITRA PHARM INC.

XX

PI Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;

XX

DR WPT; 2003-093124/08.

XX

XX New purified or isolated nucleic acids of essential genes of

PT *Aspergillus fumigatus*, useful for treating or preventing infections by

PT *A. fumigatus*, or for treating a non-infectious disease in a subject

PT e.g. cancer -

XX

PS Disclosure; Page -; 175pp; English.

XX

CC The invention relates to novel purified or isolated nucleic acids of

CC essential genes of *Aspergillus fumigatus*. The isolated nucleic acids of

CC the invention are used to treat or prevent infections by a pathogenic

CC organism such as *A. fumigatus*, to treat a non-infectious disease in a

CC subject (e.g. cancer), to prevent or contain contamination of an object

CC by *A. fumigatus*, or to prevent or inhibit formation on a surface of a

CC biofilm comprising *A. fumigatus*. The polynucleotides are useful for

CC expressing recombinant protein for characterization, screening or

CC therapeutic use, as markers for host tissues in which the pathogenic

CC organisms invade or reside, for comparing with the DNA sequence of *A.*

CC *fumigatus* to identify duplicated genes or paralogues having the same or

CC similar biochemical activity and/or function, for comparing with DNA

CC sequences of other related or distant pathogenic organisms to identify

CC potential orthologous essential or virulence genes, for selecting and

CC making oligomers for attachment to a nucleic acid array for examination

CC of expression patterns, for raising anti-protein antibodies, as an

CC antigen to raise anti-DNA antibodies or to elicit another immune

CC response, and for identifying polynucleotides encoding the other protein

CC with which binding occurs or to identify inhibitors of the binding

CC

CC interaction. The polypeptides may be used to raise antibodies or to
CC elicit immune response, as a reagent in assays designed to quantitatively
CC determine levels of the protein in biological fluids, as a marker for
CC host tissues in which pathogenic organism invade or reside, and to
CC isolate correlative receptors or ligands in the case of virulence
CC factors. This polynucleotide sequence represents one of the essential
CC genes of *Aspergillus fumigatus* of the invention.
XX
SQ Sequence 1993 BP; 543 A; 516 C; 496 G; 438 T; 0 other;

Query Match 13.2%; Score 263.8; DB 25; Length 1993;
Best Local Similarity 56.3%; Pred. No. 3.8e-51;
Matches 557; Conservative 0; Mismatches 372; Indels 60; Gaps 1;

Qy 609 ATAGTCGGTTAGGCTCCCTAGGAAGTTCTCGATCGCTTATAGGTTTGGATGCAC 668
Db 936 AGACTCTCCGACAGCCCTCGTACCTTAATAAGTACCCTATAAGTTCTCGACGCAC 995

Qy 669 CTGCTTTGCAAGATGATTTTATCTGAATCTGGTAGATGGTCTTCACACAATGTGTGG 728
Db 996 CCGATTTCGAAGATGATTTCTACTTGAATCTGGTAGATGGTCTTCACACAATGTGTGG 1055

Qy 729 CTGTTGGTTGGTAACTGTCTATCTCTGGAATCTGGAATCTGTACGACGAAGTAACTAAT 788
Db 1056 GTGTGGCTTAGCAATTCAGTGTACATGTGGAATTCGAATCCGACGGGTGACGAAC 1115

Qy 789 TATGTGATTTGGGGGTTGATGATG----- 813
Db 1116 TTTGCGAAGTAAAGATGATGATCTCACAAGTGTAGCTGATTCAGAGGGTAAATGAGAT 1175

Qy 814 -----TGTGTTGTTCTGTGTGGTCTCAAGTGTGATCTC 848
Db 1176 TTTCTTATATGTTGGCAATGATTTATGTTGTTGATGATGATTTGGACGGGCACAC 1235

Qy 849 ATCTGTGTTGGAACTAAACAATGTAAGTTCAGATTGGGATGACGACGAAGATGCAAGA 908
Db 1236 ATCTTTCAATAGAACTGGCAAGTCTTGTACAGATATGGATGCAAGCATTTGCGCC 1295

Qy 909 AGATAAGATCAATGAGGGCATCGGTTACGTGTGGGGCTTGGCTGGAGTTCAATCTC 968
Db 1296 GTCTTCGGACAATGATTTGGGCATACCAATCGTGTGGGGCTCTTGTCTTGGACGATCAT 1355

Qy 969 TTTTGTCTCTGTGGGCGGATGAAGAATATTTATCAAGAGATATACGCACACAAGAAG 1028
Db 1356 TCCTTAGCTAGGTTCTCGGATCGACTGATCTCCACCTGATGTGCTGCTCCCGATC 1415

Qy 1029 ATTTTGTAGTAACTGTGAGGACACAATCAGAGTTTGTGACTGAAGTGTCTATG 1088
Db 1416 AGTATTTACGTGCACTGTCCGGCCATTAAGCAGGAAGTTTCGGACTCCGGTGGAAACACCG 1475

Qy 1089 ATACCGTGTAGTTGGCATCTGGAGGAATGACACAAATGTTGTTGTAATCAACACT 1148
Db 1476 AAGATGGCCAACTAGCTTCAGCGGAATGACACAACTCATGTGTGGGACAGCTGA 1535

Qy 1149 CAACCCAGCTGTCTCAAGTACTGTGAGCAACAGCAGCTGTTTAAAGCTATTGCAATGGT 1208
Db 1536 ATGAGACACCGCTTTATCGCTTCTCGGACCATACCGCAGCGGTGAAGCCATCGCTGGT 1595

Qy 1209 CTCTCATCTTCATGGACTTCTTCATCTGAGGAGAACTGACAGATAGATGATTCGTT 1268
Db 1596 CACCCCATCAACACATCTCTCGCTCAGGTGGAGGACGGCGGACCGGACCATCAAT 1655

Qy 1269 TTTGGAATACACACAACTCACAACCTTAGCTGTATGGACATCGGAAGTCAAGTTTGA 1328
Db 1656 TCTGGACACTGCTACTGTTTCACTGATCAAGGAAGTTGATACGGGTAGCCAGGTGTGA 1715

Qy 1329 ATCTTGTCTGTGCAAAATGTCACGAACTAGTAGACACACATGGGTACTCCAGAAC 1388
Db 1716 ACCTGGCATGTGCGAAGAACTCGGACGAATCATCAGTACACACGGGTACAGTCAGAAC 1775

Qy 1389 AGATTATTGTTGGAGATACCCCACTATGTCAAAGCTGGCGACTCTTACCGGCCCACT 1448
Db 1776 AGATTGTAATCTGGAATACCGCGCATGGAGAGATGTTATCGCTGACCGGCCACACT 1835

Qy 1449 ATAGGTTCTTATCTTCCATCTCTCCAGATGACAGACTATTGTAACTGGAGCTGAG 1508
Db 1836 TCCGTGTTCTTACTCTGCGATGAGCCCGATGACAGACCGTGTACCGAGCTGTG 1895

Qy 1509 ATGAAGCGCTTAGGTTCTTGAATGTTTTC 1537
Db 1896 ACGAACGTTGAGATTCTGGAAGATTTTC 1924

RESULT 10
ABT17680
ID ABT17680 standard; DNA; 3398 BP.
XX
AC ABT17680;
XX
DT 16-APR-2003 (first entry)
XX
DE *Aspergillus fumigatus* essential gene #38.
XX
KW Fungicide; cytostatic; essential gene; *Aspergillus fumigatus*; infection;
KW cancer; contamination; biofilm; antibody; immune response; ds.
XX
OS *Aspergillus fumigatus*.
XX
PN WO200286090-A2.
XX
PD 31-OCT-2002.
XX
PP 23-APR-2002; 2002WO-US13142.
XX
PR 23-APR-2001; 2001US-285697P.
PR 27-APR-2001; 2001US-287066P.
PR 05-JUN-2001; 2001US-295890P.
PR 09-JUL-2001; 2001US-303899P.
PR 31-AUG-2001; 2001US-316362P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;
XX WPI; 2003-093124/08.
DR
XX
PT New purified or isolated nucleic acids of essential genes of
PT *Aspergillus fumigatus*, useful for treating or preventing infections by
PT *A. fumigatus*, or for treating a non-infectious disease in a subject
PT e.g. cancer
XX
PS Disclosure; Page -; 175pp; English.
XX
CC The invention relates to novel purified or isolated nucleic acids of
CC essential genes of *Aspergillus fumigatus*. The isolated nucleic acids of
CC the invention are used to treat or prevent infections by a pathogenic
CC organism such as *A. fumigatus*, to treat a non-infectious disease in a
CC subject (e.g. cancer), to prevent or inhibit formation of an object
CC by *A. fumigatus*, or to prevent or inhibit formation on a surface of a
CC biofilm comprising *A. fumigatus*. The polynucleotides are useful for
CC expressing recombinant protein for characterisation, screening or
CC therapeutic use, as markers for host tissues in which the pathogenic
CC organisms invade or reside, for comparing with the DNA sequence of *A.*
CC *fumigatus* to identify duplicated genes or paralogues having the same or
CC similar biochemical activity and/or function, for comparing with DNA
CC sequences of other related or distant pathogenic organisms to identify
CC potential orthologous essential or virulence genes, for selecting and
CC making oligomers for attachment to a nucleic acid array for examination
CC of expression patterns, for raising anti-protein antibodies, as an
CC antigen to raise anti-DNA antibodies or to elicit another immune
CC response, and for identifying polynucleotides encoding the other protein
CC with which binding occurs or to identify inhibitors of the binding
CC interaction. The polypeptides may be used to raise antibodies or to
CC elicit immune response, as a reagent in assays designed to quantitatively
CC determine levels of the protein in biological fluids, as a marker for
CC host tissues in which pathogenic organism invade or reside, and to

CC isolate correlative receptors or ligands in the case or virulence
CC factors. This polynucleotide sequence represents one of the essential
CC genes of Aspergillus fumigatus of the invention.
XX
SQ Sequence 3398 BP; 884 A; 845 C; 789 G; 880 T; 0 other;

Query Match	13.2%; Score 263.8; DB 25; Length 3398;
Best Local Similarity	56.3%; Pred. No. 4.5e-51;
Matches 557; Conservative	0; Mismatches 372; Indels 60; Gaps 1;
QY	609 ATAGTCGGTTAAGGCTCTAGAGGTTCCGATCGCCTTATAAGGTTTGGATGCAC 668
DB	1341 AGACTCTCGCAGCAGCCCTCGCTTAATAAAGTACCCCTATAAGGTTCTCGACGAC 1400
QY	669 CTGCTTTGCAAGATGATTTTATCTGAATCTGGTAGATTGGCTTTCACACAAATGTTGG 728
DB	1401 CCGATTTGCAAGATGATTTTCTACTTGAATCTGGTGGAGTGGAGTATGCTAG 1460
QY	729 CTGTTGGTTGGTAACTGCTGTATCTCTGAAATGCTGTAGCAGCAAGGTAACTAAT 788
DB	1461 GTGTTGGCCTAGGCAATTCAGTGTATCATGTGGAATTCGAAATCCGACGGGTGACGAAAC 1520
QY	789 TATGTGATTTGGGGTGTGATGTTG-----813
DB	1521 TTTGCGAATAGAGATGATCTGTCAAGTGTAGCTGGATTGAGGGTAAATGAGAT 1580
QY	814 -----TGTGTGTTGTTGGTGGCTCAACGTGTACTC 848
DB	1581 TTCTCTATATGTTGGCAATGATTAATATGTTTTCATGACTGACATTTGGACAGGGCACAC 1640
QY	849 ATCTTCTGCTTGGAACTCAATGTAAGTTCAGATTTGGATGCGATCAGCAAGATCAAGA 908
DB	1641 ATCTTTCAATAGGAATGGAAGGCTTGTACAGATATGGATGCAAGCATTTGTGCCC 1700
QY	909 AGATAAGATCAATGAGGGCCATCGGTTACGTTGCGGGCTTGGCGTGGAGTTTCATCTC 968
DB	1701 GTCTTCGGAATGATTTGGGCAATACCAATCGTGTGGGGCTCTTGTTCGAAACGATCATA 1760
QY	969 TTTTGTCTTCTGAGCGGATGAAGATATTTATCAAGAGATATACGCACACAGAG 1028
DB	1761 TCCTTACGTGAGTTTCTCGGATCGACTGATCTTCACCGTGTATGTTCTCCCGAGATC 1820
QY	1029 ATTTTGTAGTAACCTGTGAGACACAAATCGAGGTTTGTGGACTGAGTGGTTCATATG 1088
DB	1821 AGTATTTACGTGACTGTCGGGCCATAAGCAGGAAGTTTTCGGGACTCCGGTGGAAACCG 1880
QY	1089 ATAAACGTGATGTCATCTGGAGGAATCAACAAATGTTGTTGGATCAACACT 1148
DB	1881 AAGATGCGCAACTAGCTTCAGCGGGAATGACAACTCATGTTGTTGGGACAGCTGA 1940
QY	1149 CAACCCAGCCTGTCCTCAAGTACTGTGAGCACACAGCAGCTGTTAAAGCTATTGTCATGT 1208
DB	1941 ATGAGACACGCTTTATCGCTTCTCGACCATACCGCGCGTGAAGGCCATCGCTGCT 2000
QY	1209 CTCTCATCTTATGAGCTTCTTGATCTGGAGGAGAACTGACAGATAGATGTTATGCTT 1268
DB	2001 CACCCCATCAACACCATCTCTCGCTCAGGTGGAGGCGCGCGGACCGGACCATCAAT 2060
QY	1269 TTTTGGATCAACACCAAACTCACACTTACCTGATGACATCTGGAAGTCAAGTTTGA 1328
DB	2061 TCTGAAACATGTAATGTTTCACTGATCAGGAAGTTGATACGGTACGAGGTGTA 2120
QY	1329 ATCTTGTCTGTCCTCAAAATGTCACAACTAGTAAGCAACATGGGTACTCCGAAACC 1388
DB	2121 ACCTGGCATGCTCGAAGAACTCGGAGGAAATCATCATGATACACAGGCTACAGTCAAGAAC 2180
QY	1389 AGATTATTTGGAGTACCCCACTATGTCAAAGCTGGGACTCTTACCGGCCATCTT 1448
DB	2181 AGATTGTAATCTGGAATATCCCGGATGGAGCAGATTTGATCGCTACCGGCCACACTT 2240
QY	1449 ATAGGGTTCCTATCTTTGCCATCTCTCCAGATGGACAGACTATTTGTAACGTGGAGTGGAG 1508
DB	2241 TCCGTGTTCTTACTTTCGATGAGCCCGATGAGACCGTCTGTTTACCGGAGCTGTTG 2300

QY 1509 ATGAACGCTTAGGTTCTGGAATGTTTC 1537
DB 2301 ACGAACGTTGAGATTCGGAATGTTTC 2329

RESULT 11
ABT19494
ID ABT19494 standard; DNA; 3993 BP.
XX
XX ABT19494;
AC
AC ABT19494;
DT 16-APR-2003 (first entry)
XX
XX Aspergillus fumigatus essential gene #1852.
XX
XX Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;
XX cancer; contamination; biofilm; antibody; immune response; ds.
XX
XX Aspergillus fumigatus.
OS
XX WO200286090-A2.
FN
XX
XX
XX 31-OCT-2002.
XX
XX 23-APR-2002; 2002WO-US13142.
XX
XX 23-APR-2001; 2001US-285697P.
PR 27-APR-2001; 2001US-287066P.
PR 05-JUN-2001; 2001US-295890P.
PR 09-JUL-2001; 2001US-303899P.
PR 31-AUG-2001; 2001US-316362P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Jiang B, Tishkoff D, Zamudio C, Broshkin AM, Hu W, Lemieux SM;
XX WPI; 2003-093124/08.

New purified or isolated nucleic acids of essential genes of
Aspergillus fumigatus, useful for treating or preventing infections by
A. fumigatus, or for treating a non-infectious disease in a subject
e.g. cancer -

Disclosure; Page -; 175pp; English.

The invention relates to novel purified or isolated nucleic acids of
essential genes of Aspergillus fumigatus. The isolated nucleic acids of
the invention are used to treat or prevent infections by a pathogenic
organism such as A. fumigatus, to treat a non-infectious disease in a
subject (e.g. cancer), to prevent or contain contamination of an object
by A. fumigatus, or to prevent or inhibit formation on a surface of a
biofilm comprising A. fumigatus. The polynucleotides are useful for
expressing recombinant protein for characterisation, screening or
therapeutic use, as markers for host tissues in which the pathogenic
organisms invade or reside, for comparing with the DNA sequence of A.
fumigatus to identify duplicated genes or paralogues having the same or
similar biochemical activity and/or function, for comparing with DNA
sequences of other related or distant pathogenic organisms to identify
potential orthologous essential or virulence genes, for selecting and
making oligonucleotides for attachment to a nucleic acid array for examination
of expression patterns, for raising anti-protein antibodies, as an
antigen to raise anti-DNA antibodies or to elicit another immune
response, and for identifying polynucleotides encoding the other protein
with which binding occurs or to identify inhibitors of the binding
interaction. The polypeptides may be used to raise antibodies or to
elicit immune response, as a reagent in assays designed to quantitatively
determine levels of the protein in biological fluids, as a marker for
host tissues in which pathogenic organism invade or reside, and to
isolate correlative receptors or ligands in the case of virulence
factors. This polynucleotide sequence represents one of the essential
genes of Aspergillus fumigatus of the invention.

SQ Sequence 3993 BP; 1049 A; 991 C; 913 G; 1040 T; 0 other;

Query Match 13.2%; Score 263.8; DB 25; Length 3993;
Best Local Similarity 56.3%; Pred. No. 4.7e-51;
Matches 557; Conservative 0; Mismatches 372; Indels 60; Gaps 1;

QY 609 ATAGTCGGTTAAGGCTCCCTAGGAAGTTCTCGATCGCTTATAGGTTTGGATGCAC 668
Db 1936 AGACTCTCGCAAGCAGCTCGTACGTTAATAAAGTACCTATAAGTTTCTCGACGCAC 1995

QY 669 CTGTTTGGCAAGTATGTTTATCTGAATCTGTAGATTGGTCTTACACAATGTCTGG 728
Db 1996 CCGATTTCGAAGATGTTTCTACTTGAATCTGTGACTGGCGCAGTAGTATGTCTAG 2055

QY 729 CTGTTGTTTGGTAACTGTCTATCTCTGGAATGCTTGTAGCAGCAAGTTAACTAAT 788
Db 2056 GTGTGGCTTAGGCAATTCAGTGTACATGTGAATTCGAATACCGGACGGTGCAGAAAC 2115

QY 789 TATGTGATTTGGGGTGTGATGTTG-----813
Db 2116 TTTGCGAACTAAGAGATGATCTGTCAAGTGTAGCTGATTCAGAGGGTAAATGAGAT 2175

QY 814 -----TGTTTGTTCTGTGTTGGCTCAACGTGGTACTC 848
Db 2176 TTCTCTATATGTTGCAATGATTATGTTTTCATGACTGACATTTGGACAGGCGACAC 2235

QY 849 ATCTTGCTGTGGAACTAACTAATGTTAAAGTTTCAGATTGGGATGCGCAAGATGCAAG 908
Db 2236 ATCTTTCAATAGGAAGTGGCAAGGTTCTGTACAGATATGGGATGCGAGCATTTGCGCC 2295

QY 909 AGATAAGATCAATGAGGGCCCATCGTTTACGTGTGCGGGCTTGGCTGGAGTTCAATCTC 968
Db 2296 GTCTTGGCAATGATTGGGCATACCAATCGTGTGGGGCTTCTTGTTCGAACGATCATA 2355

QY 969 TTTTGTCTTGTGGGACGAGATGAAGATATTTATCAACGAGATATACGACACAGAAG 1028
Db 2356 TCCTTACGTAGGTTCTCGGATCGACTGATCTTCCACCGGTATGTTGCTTCCCGAGATC 2415

QY 1029 ATTTTGTAGTAACTGTCAGACACAAATCAGAGGTTTGTGACTGAAGTGTCTATG 1088
Db 2416 AGTATTACGTGACTGTTCGGCCATAGCAGGAAGTTTGGGACTCGGTGGAAACCG 2475

QY 1089 ATAACCGTGTGATCTCGAGGAATGACAAATGTTTGTGTTGTTGTTGTTGTTGTTGTTG 1148
Db 2476 AAGATGGCCAACTAGCTTTCAAGCGGAATGACAACTCATGTTGTGGGACAACTGA 2535

QY 1149 CAACCCAGCTGCTCAAGTACTGTGACACACAGCAGCTGTTAAAGCTATTTCATGGT 1208
Db 2536 ATGAGACACCGTTTATCGTCTTCGGACCATACCGAGCCGTGAGGCCATCGCTGGT 2595

QY 1209 CTCCTCATCTTCATGGAATCTTTCATCTGAGGAGGAATGCGAGATAGATGTTGTT 1268
Db 2596 CACCCCATCAACACCATCTCTCGCTCAGGTGGAGGACGCGGACCGGACCATCAAT 2655

QY 1269 TTTGGAATACACCAACCACTACACCTTAGCTGATGACACACTGGAAGTCAGGTTGCA 1328
Db 2656 TCTGGAACACTGCTACTGTTTCACTGATCAGGAAGTTGATACGGGTAGCCAGTGTGTA 2715

QY 1329 ATCTTGTCTGTGTCACAAATGTCAACGAACTAGTAAAGCAACATGGGTACTCCAGAAC 1388
Db 2716 ACCTGGCTGTGCGAAGAACTCGGACGAAATCATCAGTACACACCGCTACAGTCAGA 2775

QY 1389 AGATTATTTGTTGGAGATACCCCATCTATGTCAAGCTCGGACTCTTACCGGCGCATCTT 1448
Db 2776 AGATTGTAACTGGAATATACCGCGCATGGAGCAATGTATCGCTGACCGGCGCATCTT 2835

QY 1449 ATAGGTTCTCTATCTTGTCCATCTCTCCAGATGGACAGACTATTGTAACTGGAGTGG 1508
Db 2836 TCCGTGTTCTTACCTTGGATGAGCCCGATGGACAGACCGTCTGTTACCGGAGTGGT 2895

QY 1509 ATGAACGCTTAGGTTCTGGATGTTTTC 1537
Db 2896 ACGAAACGTTGAGATTCTGGAGATTTTC 2924

RESULT 12

ABL05492
ID ABL05492 standard; cDNA; 14283 BP.
XX ABL05492;
AC ABL05492;
XX 26-MAR-2002 (first entry)
DT
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 10958.
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX Drosophila melanogaster.
OS
XX WO200171042-A2.
PN
XX 27-SEP-2001.
PD
XX 23-MAR-2001; 2001WO-US09231.
PF
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX (PEKE) PE CORP NY.
PA Venter JC, Adams M, Li PWD, Myers EW;
PI
XX WPI; 2001-856860/75.
DR P-PSDB; ABB61389.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 10958; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 14283 BP; 4076 A; 2905 C; 2990 G; 4312 T; 0 other;

Query Match 12.9%; Score 259; DB 23; Length 14283;
Best Local Similarity 60.4%; Pred. No. 9.3e-50;
Matches 446; Conservative 0; Mismatches 290; Indels 3; Gaps 1;

QY 802 GGTGATGATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 861
Db 11771 GGAATGGAATACCGTGACCTCGGTGTCGTGGAACGCGTGCACACCGCTGCGTGGG 11830

QY 862 AACTAAACAATGTAAGTTTCAGATTGGGATGAGCAAGATGCAAGATGATCAAT 921
Db 11831 CACACATCAGGCTAGCTGACCGTCTGGGATGTTGGCGGCAATAAGCAGATCAACAAT 11890

QY 922 GGAGGGCCATCGTTACGTTACGTTGGGCGCTTGGCGGCTTATCTCTTTGTTCTTCTGG 981
Db 11891 GAATGGCCATTCGCGCGGTGTTGGCGCGCTTGGCGTGGACAGTGCATCTCTGCGAGCG 11950

QY 982 TGGACGCGGATGAAGATATTTATCAACGAGATATACGACACAGAAG---ATTTGTTAG 1038
Db 11951 GTCGCGAGACCGTTGGATCATAACAGGGGATACAGAGACGCGCAACTGCAATCGAGCG 12010

QY 1039 TAACTGTACGACACAAATCAGAGGTTTGGAGCTGAAGTGGTCAATGATGATACCGTGA 1098
Db 12011 CAGATTGCCGACATCGCGAGGAGTGTGGGACTGAAATGGTCAACGGATATCAATA 12070
QY 1099 GTTGCATCTGAGGAATGACAAATTTGTTTGGAAATCAACACTCAACCCAGCC 1158
Db 12071 CTTGCCAGTGGCGGCAACGATATCGGTTGTATGTGTGGAAATCAGCATTCGGTGAATCC 12130
QY 1159 TGTCTCAAGTACTGTGAGCACACAGCAGCTGTTAAAGCTATTGGCATGGTCTCTCATCT 1218
Db 12131 CATAATCATACACAGGACATATGGCGCTGTAAAGCGATCGCGTGTGCGCGCATCA 12190
QY 1219 TCATGCACTTCTTGCATCTGGAGAGGAACTGCAGATAGATGTTGCTTTTGGAAATAC 1278
Db 12191 CACCGATCTCTGGCAGCGCGGTGGAAACGGCGATAGGTGTATCCGTTTCTGGAATAC 12250
QY 1279 AACCAAACTCACACTTAGCTGATGAGACACATGGAAGTCAAGTTTGCATCTTGTCTG 1338
Db 12251 GCTGACGGGCCACCCATCGAGTGGTGGACACCGGCTCGCAGGTTTGCATCTCGGCTG 12310
QY 1339 GTCCAAAATGTCAAGCACTAGTAAAGCACACATGGGTACTCCAGAACCAAGATTTGT 1398
Db 12311 GTCCAAAGCACTCTCGAGCTGGTCTCCAGCACCGCTACTCGCAGAACCAAGATTTGT 12370
QY 1399 TTGAGATACCCCACTATGTCAAAGCTGGCACTCTTACCGGCCATCTTATAGGTTCT 1458
Db 12371 GTGGAATATCCCTCCCTGACCAAGTGGCCAACTGACGGCCATTCGTATCGTGTCT 12430
QY 1459 CTATCTTGGCATCTCTCAGATGGACACATTTGTAAGTGGAGCTGGAGATGAACGCT 1518
Db 12431 CTATCTGGCGCTGAGTCCCGATGGTGAAGCTATTGTTACGGGCGCGCGAGACGCT 12490
QY 1519 TAGGTTCTGGAATGTTTTC 1537
Db 12491 GCGATTTGGAACGTATTC 12509

RESULT 13

ABL18423
ID ABL18423 standard; DNA; 1369 BP.

XX ABL18423;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 6742.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical; gene; ds.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers BW;

XX WPI; 2001-656860/75.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -

XX Claim 1; SEQ ID NO 6742; 21pp + Sequence Listing; English.

XX

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL16175), expressed DNA
CC sequences (ABL161840-ABL16175) and the encoded proteins
CC (ABBS7737-ABBS72072).

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 1369 BP; 304 A; 394 C; 394 G; 277 T; 0 other;

Query Match 12.4%; Score 248; DB 23; Length 1369;

Best Local Similarity 55.9%; Pred. No. 1.5e-47;

Matches 514; Conservative 0; Mismatches 400; Indels 6; Gaps 2;

QY 636 TTCTCGATCGCCTTATAAGTTTGGATGCACTGCTTTGCAAGATGATTTTATCTGA 695

Db 396 TGCACGTGCGCCGTTACAAGATACTCGATGCGCGGAGCTGCAAGACGACTTCTACCTGA 455

QY 696 ATCTGCTAGATTGGTCTTCAACAATGTTGGCTGTGGTCTGGTAACTGTGCTATC 755

Db 456 ACCTTATGACTGGTCTGTAATAAACACCTTGGCGCTGGTCTCGGTTATC 515

QY 756 TCTGGAATGCTTTAGCAGCAAGTAACTAAATATGATTTGGGGTTGATGATGTG 815

Db 516 TCTGGAGTGGCTGCTGAGTGGTCAAGTCAAGCTATGCGACTTCAACAATAGGATAACC 575

QY 816 TTTGTTCTGTTGGTGGCTCAACGT---GGTACTCATCTGCTGTGGAACTAACAATG 872

Db 576 TGATTACAGCAGTTAGCTGGCAGCGGTGAAGCGCGGAGTGGCCATTTGCCACCGATCCG 635

QY 873 GTAAAGTTTCAAGTTTGGGATGCAAGATGCAAGAGATGAAGTCAATGGAGGGCCATC 932

Db 636 GTTACGTGACCAATTTGGGATGCGAGAACCAAAAGCAGATAAACGACTGGAAGAGCACT 695

QY 933 GGTTACGTGTCGGGCTTGGCTGAGTTCATCTCTTTTGTCTTCTGGTGGAGCGGATA 992

Db 696 CCGCGGTGTAAACGGCGCTAGCTTGGTGGCGCAACCGATTTGGCAGTGGCTCCCGGACC 755

QY 993 AGAATATTTATCAACGAGATATACGCA---CACAAAGAGATTTTGTAGTAAACTGTGAG 1049

Db 756 GCTCAATTTGACGGGGATATCGCAATCCGCGAGCAGACATACCGCTGCTTGGCG 815

QY 1050 GACAAATCAGAGTTTGGACTGAAGTGGTCAATGATGATGATGATGATGATGATGATG 1109

Db 816 GACAAAGTAGAGTGTGTGGCTTACAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 875

QY 1110 GAGGAATGACAAATTTGTTTGGTAAATCAACACTCAACCCAGCCTGTCTCAAGT 1169

Db 876 GTGGCAGGACACCGCTGTTGGTTGGACCGAGTGGCGGAGCCCATTTACGCCT 935

QY 1170 ACTGTGAGCACACAGCAGCTGTTAAAGCTATTGCAATTTGCAATTTGCTGCTGCTGCT 1229

Db 936 TCGACGAGCACAGGCGGTGTCAAAGCGCTGGGCTGTGCGGCACATAAGTCCGGTCTTC 995

QY 1230 TTGCATCTGGAGGAGGAACTGCAGATAGATGATGATGATGATGATGATGATGATGATG 1289

Db 996 TGGCCAGTGGCGGTGGGTCCCGATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1055

QY 1290 CACACCTTAGCTGTATGGACACTGGAAGTCAAGTTTGGTGGTGGTGGTGGTGGTGGTGG 1349

Db 1056 AGCTGGTGAAGTGCATCAACACCGCGCCAGATCAGAACCTGGCTGGCGCCAGGATTT 1115

QY 1350 TCAACGAACTAGTAAAGCACACATGGGTACTCCAGAACCAAGATTTGTTGGAGATACC 1409

Db 1116 CCCGAGAGCTGGTTACCAACCAACGACCGCGAGCAGCAGGTCATTTGCTTGGCGCTATC 1175

QY 1410 CCAGTATGTCAAAGCTGGCGACTCTTATCCGCGCCATCTTATAGGTTTCTCTATCTGCCA 1469

Db 1176 CTTGCTGAAGCAGATGGCCAGGCTGTCCGGGCACACACAGCGGTTACTCCACCTGTGAG 1235

QY 1470 TCTCTCCAGATGACAGACTATTGTAACTGGAGCTGGAGATGAACCGCTTAGGTTCTGGA 1529
 Db 1236 TGAGTCGGATACGAGTCCATAGTGAATGGGGCGCAGATGAGACGCTTCGCTTCTGGA 1295
 QY 1530 ATGTTTTCCCTCCCTAAA 1549
 Db 1296 CCGTTTTTACCAACAGAAA 1315

RESULT 14

ABL18422
 ID ABL18422 standard; DNA; 3369 BP.

XX ABL18422;

XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 6739.

XX Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ds.

XX Drosophila melanogaster.

OS WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell

PT interactions -

XX Claim 1; SEQ ID NO 6739; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL1840-ABL3051), expressed DNA

CC sequences (ABL01840-ABL16175) and the encoded proteins

CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pot_sequences.

XX Query Match

XX Best Local Similarity 12.4%; Score 248; DB 23; Length 3369;

XX Matches 514; Conservative 0; Mismatches 400; Indels 6; Gaps 2;

QY 636 TTCTCCATCGCCCTAATAGTTTGGATGCACTGCTTGGCAAGATGATTTTATCTGA 695

Db 1396 TGCCACGTCGCCCGTACAGATATCTGATGCGCGGAGCTGCAAGACGCTTCTACTGA 1455

QY 696 ATCTGGTAGATGTGCTTCAACAATGTGTTGGTGTGTTGGTAACTGTGTCTATC 755

Db 1456 ACCATTATGATGTGCTGCAAAACACCTGCGCGTGTCTCGGTGTCTCGTTATC 1515

QY 756 TCTGGAATGCTTTAGCAGCAAGGTAATAATTATGTGATTTGGGGGTGTGATGTG 815

Db 1516 TCTGAGTGCCTCAGTGGTCAAGTGACGCGCATGTGCGACTCAACAATGAGATAACC 1575
 QY 816 TTTGTTCTGTGTGGTGGCTCAACGT---GGTACTCATCTGTGTGTGGAACTAAACAATG 872
 Db 1576 TGATTACAGCAGTTAGCTGGCAGGTGAAGCCGCGAGTGGCCATTGGCACCAGTCCG 1635
 QY 873 GTAAAGTTTCAGATTGGGATGCGAGATGCAAGAGATAAGATCAATGAGCGGCATC 932
 Db 1636 GGTACGTGACCATTTGGGATGCGAGAACCAAAAGCAGATAAACCGACTGGAAGAGCACT 1695
 QY 933 GGTACGTGTGCGGCGCTTGGCCTTGGAGTTTCATCTCTTTTGTCTTCTGTGGTGGACGGGATA 992
 Db 1696 CGGCGGTGTAAACGGCGCTAGCTTGGTGGCGCAACCGATTGGCGAGTGGCTCCGCGACC 1755
 QY 993 AGAATTTTATCAACGAGATATAGCA---CACAAGAGATTTTGTAGTAAACTGTCAG 1049
 Db 1756 GCTCAATTTTGCAGCGGATATCCGCAATCCGCGAGCAGATACATACCCGCTGTGGCG 1815
 QY 1050 GACAAATCAGAGTTTGTGGACTGAAAGTGGTCAATATGATATAACCGTGGATGGCATCTG 1109
 Db 1816 GACAAAGCTAGAGTGTGTGGCTACAGTGGTGGCCGACCAATCGATACTTGGCCAGCG 1875
 QY 1110 GAGGAATGACAAATTTGTTTGGATCAACACTCAACCCAGCCTGTCTCAAGT 1169
 Db 1876 GTGGCAGCGCAACCGCCTGTGTGGTGGACCGACGATTGGCCGAGCCCAATTACGCCT 1935
 QY 1170 ACTGTGAGCACACAGCAGCTGTTAAAGCTATTGATGCTCTCTCATCTTTCATGGACTTC 1229
 Db 1936 TCGACGAGCACACAGCGCTGTGTCAGGCGCTGGCTGGTCCGACATAGTCCGGTCTTC 1995
 QY 1230 TTGCATCTGGAGGAGGAACTGCGAGATAGATGATTCGTTTTTGGAAATACAAACCAAACT 1289
 Db 1996 TGGCCAGTGGCGTGGGTCCGCGATCGCTGTGGCTTTTGGAACTGCTCACCGGGA 2055
 QY 1290 CACACCTTAGCTGTATGACACTGGAAGTCAAGTTTGCATCTTGTCTGTCCAAAATG 1349
 Db 2056 AGCTGGTGAAGTGCATCAACACCGCGCCAGATCAGCACTGGCTGGCGCCAGGATT 2115
 QY 1350 TCACGAACTAGTAAGCACATGGGTACTCCCGAAGAACCAAGATTTGTTTGGAGATACC 1409
 Db 2116 CCCGAGAGCTGGTTACCAACCCACGACACGCGAGCCACAGGTCAATTGCGCTGGCGCTATC 2175
 QY 1410 CCATATGTCAAAGCTGGCACTCTTACCGGCCATCTTATAGGGTTCTCTATCTTGCCA 1469
 Db 2176 CTTGCTGAAGCAGATGGCCAGGCTGTCCGGGCACACAGCGGTTACTCCACCTGTGAG 2235
 QY 1470 TCTCTCCAGATGGACAGACTATTGTAACTGGAGCTGGAGATGAACCGCTTAGTGTGGA 1529
 Db 2236 TGAGTCCGGATAACGAGTCCATAGTGAATGCTGGCGCGCAGATGAGACGCTTCGCTTCTGGA 2295
 QY 1530 ATGTTTTCCCTCCCTAAA 1549
 Db 2296 CCGTTTTTACCAACAGAAA 2315

RESULT 15

ABL05448

ID ABL05448 standard; cDNA; 5058 BP.

XX AC ABL05448;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 10826.

XX KW Drosophila; developmental biology; cell signalling; insecticide;

XX KW pharmaceutical; gene; ss.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

PD 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX P-PSDB; ABB61345.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Claim 1; SEQ ID NO 10826; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABB57737-ABB72072).
XX
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 5058 BP; 1244 A; 1301 C; 1304 G; 1209 T; 0 other;
Query Match 12.4%; Score 248; DB 23; Length 5058;
Best Local Similarity 55.9%; Pred. No. 2.3e-47;
Matches 514; Conservative 0; Mismatches 400; Indels 6; Gaps 2;
QY 636 TTCTCGATCGCCTATTAAGTTTGGATGCACTGCTTGGCAAGATGATTTTATCTGA 695
DB 2667 TGGCAGCTGCGCGTACAGATATCTCGATGCGCGAGCTGCAAGACGACTTCACTGA 2726
QY 696 ATCTGTAGATGCTTTCACAGATGTTGCTGTTGGTTGGTTGGTTGGTTGGTTGGTTGG 755
DB 2727 ACCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2786
QY 756 TCTGGAATGCTTGTAGCAGCAAGTAACTAAATTTATGATTTGGGGTTGATGATGTTG 815
DB 2787 TCTGAGTGGCGTCAAGTGGTCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2846
QY 816 TTGTTCTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 872
DB 2847 TGATTACAGAGTTAGTGTGCGACGGTGAAGCGCGGCGGCGGCGGCGGCGGCGGCGG 2906
QY 873 GTAAAGTTTCAGATTTGGGATGCAAGATGCAAGAGATGAAGATCAATGAGGGGCATC 932
DB 2907 GGTACGTGACCAATTTGGATGCGGAGAACCAAGACAGATGAACCGACTGGAAGACT 2966
QY 933 GGTACGTGCGGGCGCTTGGCTGGAGTTCACTCTTTTGTCTTCTGTTGGAGCGGATA 992
DB 2967 CGGCGCGTGAACGCGCTAGCTTGGTGGCGCAACCGATTTGGCGAGTGGCTCCCGGACC 3026
QY 993 AGATATTTATCAACAGATATACGCA---CACAAGAGATTTTGTAGTAAACTGTCTAG 1049
DB 3027 GCTCAATTTTCACGGGGATATCCGCAATCGCGCGAGCAGATAAACCGCTGCTGGCG 3086
QY 1050 GACAAATCAGAGTTTGTGACTGAAGTGGTCAATGATGAACCGTGAAGTTGGCATCTG 1109
DB 3087 GACAAAGCTAGAGTGTGGCTCAGATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3146
QY 1110 GAGGAATGCAACAATTTGTTTGGATCAACACTCAACCCAGCCCTGCTCTCAAGT 1169
DB 3147 GTGGAGCGCAACACCGCTGTGTTGTTGGACCGAGATGGCGGAGCCCATTTACGCT 3206

QY 1170 ACTGTAGCACACAGCAGCTGTTAAAGCTATTGAGTGGTCTCTCTCATCTTCATGAGCTTC 1229
DB 3207 TCGAGGACACAAAGCGGTGGTCAAGCGCTGGTGGCTGGTGGCTGGTGGCTGGTGGCTTC 3266
QY 1230 TTGCATCTGGAGGAGGAACATGCAATAGATGATTTGTTTGGTGAATACAAACCAAACT 1289
DB 3267 TGGCCAGTGGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3326
QY 1290 CACACCTTAGCTGTATGGACACTGGAAGTCAAGTTTGCATCTTGTCTGGTCCAAATATG 1349
DB 3327 AGCTGGTGAAGTGCATCAACACCGGCGCCAGATCAACCACTGGCTGGGCGCAGGATTT 3386
QY 1350 TCAACGAACCTAGTAAAGCACACATGGGTACTCCACAGAACCCAGATTTATTTGGAGATACC 1409
DB 3387 CCCGAGAGCTGGTTACCAACCCACGACACGCGCAGCCACAGGTCTATTGGCTGGCGCTATC 3446
QY 1410 CCACATGTCTAAAGCTGGCGACTCTTACCGGCCATCTATTATAGGTTCTCTATCTTGCCA 1469
DB 3447 CCTCGCTGAAGCAGATGGCCAGGCTGTCGCGGCACACACAGCGGTACTCCACTGTCTAG 3506
QY 1470 TCTCTCCAGATGGACAGACTATTGTAACTGGAGCTGGAGATGAAACGCTTAGGTTCTGGA 1529
DB 3507 TGAGTCCGGATTAACGAGTCCATAGTACTGGCGGCGCAGATGAGACGCTTCGCTTCTGGA 3566
QY 1530 ATGTTTCCCTTCCCTCCCTAAA 1549
DB 3567 CCGTTTTTACCACACAGAAA 3586

Search completed: January 23, 2004, 17:48:35
Job time : 603 secs

QY 1037 AGTAACTGTGAGGACAAATCAGAGGTTTGTGGACTGAAGTGGTCTATATGATAACCGT 1096
 Db 1099 GCCACATGAGTGCCATAGCCAGGAGTATGTGGCTGGCTGGCCCGCCAGATGGACGA 1158
 QY 1097 GAGTTGCACTCTGGAGGAATGACAAATGTTTGTGGTGGTGGTGGTGGTGGTGGTGGT 1144
 Db 1159 CATCTGCAAGCGGTGGCAATGATAATGTTCAAGTGTGGCTGGTGGTGGTGGTGGTGGT 1218
 QY 1145 CACTCAACCCAGCTGCTCAAGTACTGTGAGCAGACAGCAGCTGTTAAAGCTATTGCA 1204
 Db 1219 AGTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1278
 QY 1205 TGGTCTCTCATCTTCAATGAGCTTTTGGCATCTGGAGGAGAACTGCGAGATAGATATT 1264
 Db 1279 TGGTGTCCCTGGCAGTCCAATATCTGGCAACAGGAGGAGTACCAGTGCACGACAT 1338
 QY 1265 CGTTTTGGATACAAACACCAAACTCACCTTAGCTGTATGACACAGTGGAGTTCAGGTT 1324
 Db 1339 CGCAATTTGGAGCTGCTCTGGAGCTGTCTGAGTCTGTGATGTGCTGATGTGCTGAT 1398
 QY 1325 TGCAATCTTGTCTGGTCCAAATATGTCAGCAAACTAGTAAAGCAGACATGGGTACTCCCAG 1384
 Db 1399 TGTCCATCTCTGCTCTGCCACTATAAGGAGCTCATCTCAGGCCATGGCTTTGCCAG 1458
 QY 1385 AACAGATATTGTTGGAGTACCCCACTATGTCAGAGCTGGGACTCTTACCGGCCAT 1444
 Db 1459 AACAGCTGGTATTGGAGTACCCCACTATGTCAGAGCTGGGACTCTTACCGGCCAT 1518
 QY 1445 ACTTATAGGTTCTCTATCTTGTCCATCTCTCCAGATGGAGACATATTGTAAGTGGAGCT 1504
 Db 1519 ACAGCCGGTCTCTGAGTCTCACCATGATGATCCAGAGCGGGCCAGTGGCATCTGCAGCA 1578
 QY 1505 GGAGATGAAAGCTTAGTCTGG 1528
 Db 1579 GCCGATGAGACTCTGGGCTCTGG 1602

RESULT 2

PCT-US95-01806-1
 ; Sequence 1, Application PC/TUS9501806
 ; GENERAL INFORMATION:
 ; APPLICANT: Weinstein, Jasminder
 ; TITLE OF INVENTION: Novel Mammalian Cell Cycle Protein
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Amgen Inc./Patent Operations/RW
 ; STREET: 1840 DeWittland Drive
 ; CITY: Thousand Oaks
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 91320-1789
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US95/01806
 ; FILING DATE:
 ; CLASSIFICATION:
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1767 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: unknown
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 193..1692
 PCT-US95-01806-1

Query Match

10.0%; Score 200.4; DB 5; Length 1767;

Best Local Similarity 52.8%; Pred. No. 9.7e-42;
 Matches 488; Conservative 0; Mismatches 421; Indels 15; Gaps 2;
 QY 620 AAGCTCTCTAGGAAGTTCTCTCGATCGCTTATAAGGTTTGGATGACCTGCTTTGCAA 679
 Db 679 AAGCTCTGAGATACATCTCTCTCCGCGAGAGGATTTCTGATGCCCTGAAATCGG 738
 QY 680 GATGATTTTATCGAATCTGGTAGATGGTCTTTCACAAATGTTGGCTGTGTTGGTTG 739
 Db 739 AATGACTACTACCTGAATCTTGTGATTTGGAGCTCTGGAATATGATTAGCTGTGGCACTG 798
 QY 740 GGTAACTGTGCTATCTCTGGAATGCTTGTAGCAGCAAGGTAACTAAATATTGATGATTG 799
 Db 799 GACAACTGTGCTTACTTATGGAACGCTGTTCCGCTGACATCTCTGAGCTGTTGCAATG 858
 QY 800 G---GGGTTGATGATGTTGTTGTTCTGTTGGTGGCTCAAAGTGTACTCATCTTGTCT 856
 Db 859 GAGCAGCTGGGAGTACATATCATCTCGTGGCTTGAATCAAAGAGGGCACTACTCGCT 918
 QY 857 GTTGAACCTAAATGTTAAAGTTTCAGATTTGGATTCAGCAGATGCAAGATCAAGATAAGA 916
 Db 919 GTGGGACACGATGATGCTGAGTGCAGCTATGGAATGTGAGCAGCAGAGAACGCTTGA 978
 QY 917 TCAATGAGGGCCATCGTTTACGTGTGGGCTTGGCTGGAGTTCACTCTCTTTGTCT 976
 Db 979 AACATGACCGACCACTCTGCTCGAGTAAGCTCCCTGAGTTGGAACAGCTATATCTGTCA 1038
 QY 977 TCTGTGAGCGGATAAGAAATTTATCAAGAGATATACGACACAAAGAGATTTTGT 1036
 Db 1039 AGTGTGTTCAAGATCTGGCCATCCACCACCATGTTTCGATGAGAGAACCAATGTTG 1098
 QY 1037 AGTAACTGTGAGGACACAAATCAGAGGTTTGTGGACTGTGAAGTGGTCAATGATAACCGT 1096
 Db 1099 GCCACACTGAGTGGCCATAGCCAGGAAGTATGTGGCTGCGCTGGGCCCCAGATGAGCA 1158
 QY 1097 GAGTTGGCATCTGGAGGAATGACACAAATGTTTGTGGTGGTGGTGGTGGTGGTGGTGG 1144
 Db 1159 CATCTGGCAAGCGGTGGCAATGATTAACATGTCAACGTTGGCTTGGTGGTGGTGGTGG 1218
 QY 1145 CACTCAACCCAGCTGTCTCAAAGTACTGTGAGCAGACAGCAGCTGTTTAAAGCTATTGCA 1204
 Db 1219 AGTGGCTGGTTCCTCTGAGACATTTCACTCAACATCAAGGTGCTGTCAAGGCTGTTGCA 1278
 QY 1205 TGGTCTCTCATCTTTCAGTGGCTTCTTGCATCTGGAGGAGAACTGCGAGATAGATGATT 1264
 Db 1279 TGGTGTCTCTGGCAGTCCAATATCTGGCAACAGGAGAGGTACAGTGAACGACACAT 1338
 QY 1265 CGTTTTTGGATACAAACCAACAACTCAACCTTACTGTATGACACATCTGGAAGTCAAGTT 1324
 Db 1339 CGCAATTTGGAACGCTGCTCTGAGGCTGTCTGAGTGTCTGATGTGCTATCTCCAGGTG 1398
 QY 1325 TGCAATCTTGTCTGTGTCAAAATGTCACGAACTAGTAAAGCAGACACATGAGTGGTCTCCAG 1384
 Db 1399 TGTCTCATCTCTGGTCTCTCCCACTATAAGGAGCTCATCTCAGGCCATGGCTTTGGCCAG 1458
 QY 1385 ACCAGATATTGTTGGAGTACCCCACTATGTCAGAGCTGGGACTCTTACCGGCCAT 1444
 Db 1459 AACAGCTGGTATTGAGTACCCCACTATGTCAGAGGTTGGCAGAGCTCAAGGTCAAC 1518
 QY 1445 ACTTATAGGTTCTCTATCTTGGCATCTCTCCAGATGAGCAGATATTGTAAGTGGAGCT 1504
 Db 1519 ACAGCCGGTCTCTGAGTCTCACCATGATGATCCAGAGCGGGCCAGTGGCATCTGCAGCA 1578
 QY 1505 GGAGATGAAAGCTTAGTCTGG 1528
 Db 1579 GCCGATGAGACTCTGGGCTCTGG 1602

RESULT 3

US-08-988-856B-3
 ; Sequence 3, Application US/08988856B
 ; Patent No. 6291642
 ; GENERAL INFORMATION:

APPLICANT: Weinstein, Jasminder
TITLE OF INVENTION: No. 6291642el Mammalian Cell Cycle Protein
FILE REFERENCE: 06843.0026-04; A-283 D
CURRENT APPLICATION NUMBER: US/08/988,856B
CURRENT FILING DATE: 1997-12-11
NUMBER OF SEQ ID NOS: 31
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 1700
TYPE: DNA
ORGANISM: Homo sapiens
US-08-988-856B-3

Query Match 9.9%; Score 198.8; DB 3; Length 1700;
Best Local Similarity 52.7%; Pred. No. 2.4e-41;
Matches 487; Conservative 0; Mismatches 422; Indels 15; Gaps 2;

QY 620 AAGGCTCTAGAGGTTCTCTGATCGCTTAAAGTTTGGATGCACTGCTTTGCAA 679
DB 603 AAGACCTGCGGTACATTCCTCCCTGCCAGACCGCTATCTGGATGCGCTGAATCCGA 662
QY 680 GATGATTTTATCTGAATCTGTGATGCTTCTTACACAAATGTTGGCTGTTGTTG 739
DB 663 AATGACTATTAACCTTGATGATGAGTTCTGGGATGTAAGTGGCTGCACTG 722
QY 740 GGTAACTGTCTATCTCTGGAATGCTTGTAGCAGCAAGTAACATAATATGTAATTG 799
DB 723 GACCAACAGTGTGTACCTGTGGAGTGAAGCTCTGTGATCATCTCTGAGCTTTTGC 782
QY 800 G---GGGTGATGATGTTGTTCTGTTGTTGGCTCAACGCTGCTATCTTCT 856
DB 783 GAGCAGCTGCGGTACATTCCTTCCCTGCCAGACCGTATCTGGATGCGCTGAATCCGA 662
QY 857 GTTGGAACCTAACTAGTAAAGTTTCAAGTTTGGATGCAAGATGCAAGAAAGTAAGA 916
DB 843 GTGGGCAACAGCAGTGTGAGTGCAGCTATGGGATGTGAGGAGGCAACTACTTGCT 902
QY 917 TCATGAGGAGGCGCATCGTTACGTTAGTGTGGGCTTGGCTGAGTTCTCTTTTGTCT 976
DB 903 AATATGACCACTGCTCTGCGGCTGCTTAAAGCTGGAACAGCTATATCTCTGCTCC 962
QY 977 TCTGGTGGAGCGGATAAGATAATTTATCAACGAGATATACGCAACAAGAAATTTGTT 1036
DB 963 AGTGGTTTACGTTCTGGCCACATCCACCACCATGATGTTGGGTAGCAGAACCATGTG 1022
QY 1037 AGTAACTGTGAGACACAAATCAGAGGTTTGTGATGAGTGTGATATGATACCTG 1096
DB 1023 GCCACACTGAGTGGCCACAGCAGGAGTGTGTGGCTGCTGGGCTGAGTGAAGCA 1082
QY 1097 GAGTTGGCATCTGGAGGAATGACAAATAATGTTGTTGTAATCAACACTCAACCCAG 1156
DB 1083 CATTTGGCCAGTGTGTGATGATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1142
QY 1157 -----CCTGTCTCAAGTACTGTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1204
DB 1143 GGTGGCTGGGTTCTCTGAGACATTCACCCAGCATCAGGGGCTGTCAAGGCGGTAGCA 1202
QY 1205 TGGTCTCTCATCTTTCATGAGCTTCTGTCATCTGGAGGAGGAACTGCAAGATGATATT 1264
DB 1203 TGGTGTCTCTGGAGTCCATGCTCTGGCAACAGAGGAGGAGGAGGAGGAGGAGGAGG 1262
QY 1265 CGTTTTTGGAAATACAAACACAACTCAACCTTAGCTGTATGAGCAGCTGGAAGTCAAGTT 1324
DB 1263 CGCATCTGGAATGTGCTCTGCGGCTGTGCTGAGTGGCTGAGTGGCATTCCTCCAGGTG 1322
QY 1325 TGAATCTTGTCTGTTCCAAATATGTCAGCACTAGTAGCAGCAGCAGCAGCAGCAGCAG 1384
DB 1323 TGCTCCATCTCTGTTCTGCCATTACAGAGCTCATCTCAGGCCATGCTTTCACAG 1382
QY 1385 AACCAAGATTATTTTGGAGATACCCCACTATGTCAAAGCTGGCGACTCTTACCGGCCAT 1444
DB 1383 AACCAAGTAGTTATTTGGAGTACCCCAACCATGTCGCAAGGTGCTGAATCAAGAGTCA 1442

QY 1445 ACTTATAGGGTCTCTATCTTCCATCTCTCCAGATGAGCAGACTATTGTAACTGAGCT 1504
DB 1443 ACATCCCGGTCTGTAGTCTGACCATGAGCCAGATGGGCCACACAGTGGCATCCGAGCA 1502
QY 1505 GGAGATGAACAGCTTAGGTTCTGG 1528
DB 1503 GCAGATGAGACCTGAGGCTATGG 1526

RESULT 4
PCT-US95-01806-3
Sequence 3, Application PC/TUS9501806
GENERAL INFORMATION:
APPLICANT: Weinstein, Jasminder
TITLE OF INVENTION: Novel Mammalian Cell Cycle Protein
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Angen Inc./Patent Operations/RBW
STREET: 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/01806
FILING DATE:
CLASSIFICATION:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1700 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 117..1616
PCT-US95-01806-3

Query Match 9.9%; Score 198.8; DB 5; Length 1700;
Best Local Similarity 52.7%; Pred. No. 2.4e-41;
Matches 487; Conservative 0; Mismatches 422; Indels 15; Gaps 2;

QY 620 AAGGCTCTAGAGGTTCTCTGATCGCTTAAAGTTTGGATGCACTGCTTTGCAA 679
DB 603 AAGACCTGCGGTACATTCCTTCCCTGCCAGACCGTATCTGGATGCGCTGAATCCGA 662
QY 680 GATGATTTTATCTGAATCTGTGATGCTTCTTACACAAATGTTGGCTGTTGTTG 739
DB 663 AATGACTATTAACCTTGATGATGAGTTCTGGGATGTGAGGAGGCAACTACTTGCT 722
QY 740 GGTAACTGTCTATCTCTGGAATGCTTGTAGCAGCAAGTAACATAATATGTAATTG 799
DB 723 GACCAACAGTGTGTACCTGTGGAGTGAAGCTCTGTGATCATCTCTGAGCTTTTGC 782
QY 800 G---GGGTGATGATGTTGTTGTTGTTGGCTCAACGCTGCTATCTTCTTGTCT 856
DB 783 GAGCAGCTGCGGTACATTCCTTCCCTGCCAGACCGTATCTGGATGCGCTGAATCCGA 842
QY 857 GTTGGAACCTAACTAGTAAAGTTTCAAGTTTGGATGCAAGATGCAAGAAAGTAAGA 916
DB 843 GTGGGCAACAGCAGTGTGAGTGCAGCTATGGGATGTGAGGAGGCAACTACTTGCT 902
QY 917 TCATGAGGAGGCGCATCGTTACGTTAGTGTGGGCTTGGCTGAGTTCTCTTTTGTCT 976
DB 903 AATATGACCACTGCTCTGCGGCTGCTTAAAGCTGGAACAGCTATATCTCTGCTCC 962
QY 977 TCTGGTGGAGCGGATAAGATAATTTATCAACGAGATATACGCAACAAGAAATTTGTT 1036

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/204,675
FILING DATE: 02-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.45501
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202.508.9100
TELEFAX: 202.508.9299
TELEX: 197430 BMB UT
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4164 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: PAT-1
US-08-204-675-1

Query Match 2.3%; Score 46.4; DB 1; Length 4164;
Best Local Similarity 53.3%; Pred. No. 0.035;
Matches 98; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 1822 CTCATTTTATGAAAGATGATACAAATGGTAAATTTATGCTTGGACTTATACAT 1881
Db 3841 CTTATCGTTAATCGAATGATATCTATTAACTGCTTTCTGCTAATAATATATAT 3900
QY 1882 GCATTGATGAGTGTAGCCAAAGTTTTTTTATTTACTCTTTTCTCTCTTTTGAT 1941
Db 3901 GTAAAGTACGCTTTTGTGGAATTTTTTAAACCTTGTTATTTTTTCTTCATCC 3960
QY 1942 AGTGCTCTCGCTGATTTATATATTTTAAAGTGGTTACAGAGAAAAA 2001
Db 3961 GTAACTCTTACCTCTTTTATTTACTTTCTAAATCCAAATCAAAACATATAA 4020
QY 2002 AAAA 2005
Db 4021 TAAA 4024

RESULT 8
US-08-660-754-1
Sequence 1, Application US/08660754
Patent No. 5843772
GENERAL INFORMATION:
APPLICANT: Devine, Scott E.
APPLICANT: Boeke, Jef D.
APPLICANT: Braiterman, Lelita T.
TITLE OF INVENTION: In Vitro Transposition of Artificial
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner, Birch, McKie, and Beckett
STREET: 1001 G Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/660,754
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/204,675
FILING DATE: 02-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.45501
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202.508.9100
TELEFAX: 202.508.9299
TELEX: 197430 BMB UT
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4164 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: PAT-1
US-08-660-754-1

Query Match 2.3%; Score 46.4; DB 2; Length 4164;
Best Local Similarity 53.3%; Pred. No. 0.035;
Matches 98; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 1822 CTCATTTTATGAAAGATGATACAAATGGTAAATTTATGCTTGGACTTATACAT 1881
Db 3841 CTTATCGTTAATCGAATGATATCTATTAACTGCTTTCTGCTAATAATATATAT 3900
QY 1882 GCATTGATGAGTGTAGCCAAAGTTTTTTTATTTACTCTTTTCTCTCTTTTGAT 1941
Db 3901 GTAAAGTACGCTTTTGTGGAATTTTTTAAACCTTGTTATTTTTTCTTCATCC 3960
QY 1942 AGTGCTCTCGCTGATTTATATATTTTAAAGTGGTTACAGAGAAAAA 2001
Db 3961 GTAACTCTTACCTCTTTTATTTACTTTCTAAATCCAAATCAAAACATATAA 4020
QY 2002 AAAA 2005
Db 4021 TAAA 4024

RESULT 9
US-08-796-364-1
Sequence 1, Application US/08796364
Patent No. 5968785
GENERAL INFORMATION:
APPLICANT: Devine, Scott E.
APPLICANT: Boeke, Jef D.
APPLICANT: Braiterman, Lelita T.
TITLE OF INVENTION: In Vitro Transposition of Artificial
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner, Birch, McKie, and Beckett
STREET: 1001 G Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/796,364
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/204,675
 FILING DATE: 02-MAR-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Kagan, Sarah A.
 REGISTRATION NUMBER: 32,141
 REFERENCE/DOCKET NUMBER: 01107.45501
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202.508.9100
 TELEFAX: 202.508.9299
 TELEX: 197430 BMB UT
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 4164 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: circular
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 IMMEDIATE SOURCE:
 CLONE: PAT-1
 US-08-796-364-1

Query Match 2.3%; Score 46.4; DB 2; Length 4164;
 Best Local Similarity 53.3%; Pred. No. 0.035;
 Matches 98; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
 QY 1822 CTCATTTTATATGAAAGATGATAACAAATGGGTAATTTATTGTCCTGGACTTATACAT 1881
 DB 3841 CTTATCGTTAATCGAATGATATCTATTAACTGCTTTTCTGTCTAAATAATATAT 3900
 QY 1882 GCATTGAGGAGTGTAGCCAAAGTTTTTTTATTACTCTTTTCTTTCTTTTTCAT 1941
 DB 3901 GTAAAGTAGCGCTTTTGTGTAATTTTAAACCTTTGTTTATTATTTTCTTCATCC 3960
 QY 1942 AGTGTCTCTCCGCAATTTTATATAATTTTAAAGATGGGTTAAACAGAGAAAAA 2001
 DB 3961 GTAACTCTCTACCTCTTTTATTACTTTTCTTAAATCCAAATACAAATATAA 4020
 QY 2002 AAAA 2005
 DB 4021 TAAA 4024

RESULT 10
 PCT-US95-02520-1
 Sequence 1, Application PC/TUS9502520
 GENERAL INFORMATION:
 APPLICANT: The Johns Hopkins University
 TITLE OF INVENTION: In Vitro Transposition of Artificial
 TITLE OF INVENTION: Transposons
 NUMBER OF SEQUENCES: 7
 CORRESPONDENCE ADDRESS:
 ADDRESSES: Banner, Birch, McKie, and Beckett
 STREET: 1001 G Street, N.W.
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20001
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/02520
 FILING DATE: 02-MAR-1995
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:

NAME: Kagan, Sarah A.
 REGISTRATION NUMBER: 32,141
 REFERENCE/DOCKET NUMBER: 01107.49245
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202.508.9100
 TELEFAX: 202.508.9299
 TELEX: 197430 BMB UT
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 4164 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: circular
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 IMMEDIATE SOURCE:
 CLONE: PAT-1
 PCT-US95-02520-1
 Query Match 2.3%; Score 46.4; DB 5; Length 4164;
 Best Local Similarity 53.3%; Pred. No. 0.035;
 Matches 98; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
 QY 1822 CTCATTTTATATGAAAGATGATAACAAATGGGTAATTTATTGTCCTGGACTTATACAT 1881
 DB 3841 CTTATCGTTAATCGAATGATATCTATTAACTGCTTTTCTGTCTAAATAATATAT 3900
 QY 1882 GCATTGAGGAGTGTAGCCAAAGTTTTTTTATTACTCTTTTCTTTCTTTTTCAT 1941
 DB 3901 GTAAAGTAGCGCTTTTGTGTAATTTTAAACCTTTGTTTATTATTTTCTTCATCC 3960
 QY 1942 AGTGTCTCTCCGCAATTTTATATAATTTTAAAGATGGGTTAAACAGAGAAAAA 2001
 DB 3961 GTAACTCTCTACCTCTTTTATTACTTTTCTTAAATCCAAATACAAATATAA 4020
 QY 2002 AAAA 2005
 DB 4021 TAAA 4024

RESULT 11
 US-08-204-675-2
 Sequence 2, Application US/08204675
 Patent No. 567170
 GENERAL INFORMATION:
 APPLICANT: Devine, Scott E.
 APPLICANT: Boeke, Jef D.
 APPLICANT: Braiterman, Lelita T.
 TITLE OF INVENTION: In Vitro Transposition of Artificial
 TITLE OF INVENTION: Transposons
 NUMBER OF SEQUENCES: 7
 CORRESPONDENCE ADDRESS:
 ADDRESSES: Banner, Birch, McKie, and Beckett
 STREET: 1001 G Street, N.W.
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20001
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/204,675
 FILING DATE: 02-MAR-1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Kagan, Sarah A.
 REGISTRATION NUMBER: 32,141
 REFERENCE/DOCKET NUMBER: 01107.45501
 TELECOMMUNICATION INFORMATION:

TELEX: 197430 BBMB UT
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 4933 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: PAT-2
US-08-796-364-2

Query Match 2.3%; Score 46.4; DB 2; Length 4933;
Best Local Similarity 53.3%; Pred. No. 0.037;
Matches 98; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 1822 CTCATTTTATGAAAGATGATACAAATGGTAAATTTATGCTTGGACTTATACAT 1881
Db 4610 CTTATCGTTAATCGAATGTATATCTATTAATCTGCTTTTCTGTCTAATAATATATAT 4669

QY 1882 GCATTGATGGAGTGTAGCCAAAGTTTTTATTAATCTCTTTTCTTCTTTTGTAT 1941
Db 4670 GTAAAGTACGCTTTTGTGTAATTTTAAACCTTTGTTATTTTCTTCTTCTTCC 4729

QY 1942 AGTGCTCTCTGCAATTTATATATTAATTTAAGATGGTTAAACAGAGAAAAA 2001
Db 4730 GTAACTCTCTACCTTCTTATTTTACTTTTCTTAAATCCAAATCAAAACATAAAATAAA 4789

QY 2002 AAAA 2005
Db 4790 TAAA 4793

RESULT 14
PCT-US95-02520-2
Sequence 2, Application PC/TUS9502520
GENERAL INFORMATION:
APPLICANT: The Johns Hopkins University
TITLE OF INVENTION: In Vitro Transposition of Artificial
TITLE OF INVENTION: Transposons
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner, Birch, McKie, and Beckett
STREET: 1001 G Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02520
FILING DATE: 02-MAR-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.49245
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202.508.9100
TELEFAX: 202.508.9299
TELEX: 197430 BBMB UT
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 4933 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular

MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: PAT-2
PCT-US95-02520-2

Query Match 2.3%; Score 46.4; DB 5; Length 4933;
Best Local Similarity 53.3%; Pred. No. 0.037;
Matches 98; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 1822 CTCATTTTATGAAAGATGATACAAATGGTAAATTTATGCTTGGACTTATACAT 1881
Db 4610 CTTATCGTTAATCGAATGTATATCTATTAATCTGCTTTTCTGTCTAATAATATATAT 4669

QY 1882 GCATTGATGGAGTGTAGCCAAAGTTTTTATTAATCTCTTTTCTTCTTTTGTAT 1941
Db 4670 GTAAAGTACGCTTTTGTGTAATTTTAAACCTTTGTTATTTTCTTCTTCTTCC 4729

QY 1942 AGTGCTCTCTGCAATTTATATATTAATTTAAGATGGTTAAACAGAGAAAAA 2001
Db 4730 GTAACTCTCTACCTTCTTATTTTACTTTTCTTAAATCCAAATCAAAACATAAAATAAA 4789

QY 2002 AAAA 2005
Db 4790 TAAA 4793

RESULT 15
US-09-890-848-16
Sequence 16, Application US/09890848
Patent No. 6573053
GENERAL INFORMATION:
APPLICANT: Birth, Greg
APPLICANT: Odedra, Rajesh Muru
TITLE OF INVENTION: Analysis Method
FILE REFERENCE: PA9903
CURRENT APPLICATION NUMBER: US/09/890,848
CURRENT FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: PCT/GB00/00346
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: EP 99300873.9
PRIOR FILING DATE: 1999-02-05
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patent in Ver. 2.1
SEQ ID NO 16
LENGTH: 5277
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: plasmid
US-09-890-848-16

Query Match 2.3%; Score 46.4; DB 4; Length 5277;
Best Local Similarity 53.3%; Pred. No. 0.038;
Matches 98; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 1822 CTCATTTTATGAAAGATGATACAAATGGTAAATTTATGCTTGGACTTATACAT 1881
Db 4954 CTTATCGTTAATCGAATGTATATCTATTAATCTGCTTTTCTGTCTAATAATATATAT 5013

QY 1882 GCATTGATGGAGTGTAGCCAAAGTTTTTATTAATCTCTTTTCTTCTTTTGTAT 1941
Db 5014 GTAAAGTACGCTTTTGTGTAATTTTAAACCTTTGTTATTTTCTTCTTCTTCC 5073

QY 1942 AGTGCTCTCTGCAATTTATATATTAATTTAAGATGGTTAAACAGAGAAAAA 2001
Db 5074 GTAACTCTCTACCTTCTTATTTTACTTTTCTTAAATCCAAATCAAAACATAAAATAAA 5133

QY 2002 AAAA 2005
Db 5134 TAAA 5137

us-09-701-572-1.rni

Sun Jan 25 17:25:16 2004

Search completed: January 23, 2004, 21:08:23
Job time : 142 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 23, 2004, 19:53:14 ; Search time 713 Seconds
(without alignments)
10062.272 Million cell updates/sec

Title: US-09-701-572-1
Perfect score: 2006
Sequence: 1 gattcgacaggaagaaaa.....agaaaaaa 2006

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2356869 seqs, 1788235258 residues

Total number of hits satisfying chosen parameters: 4713738

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US05_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	326.2	16.3	1143	12	US-10-369-493-27884
2	303.8	15.1	1794	13	Sequence 27884, A
3	277	13.8	1312	12	Sequence 6595, Ap
4	277	13.8	1312	12	Sequence 26005, A
5	277	13.8	1312	12	Sequence 26011, A
6	264.6	13.2	1116	12	Sequence 26013, A
7	263.8	13.2	1398	15	Sequence 3626, A
8	263.8	13.2	1398	15	Sequence 1038, Ap
9	263.8	13.2	1875	15	Sequence 2038, Ap
10	263.8	13.2	1993	15	Sequence 7038, Ap
11	263.8	13.2	3398	15	Sequence 6038, Ap
12	263.8	13.2	3593	15	Sequence 38, Appl
13	225.6	11.2	1701	12	Sequence 5038, Ap
14	213.6	10.6	1344	10	Sequence 45642, A
15	213.6	10.6	1344	12	Sequence 988, App

ALIGNMENTS

RESULT 1

US-10-369-493-27884
; Sequence 27884, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 27884
; LENGTH: 1143
; TYPE: DNA
; ORGANISM: Neurospora crassa
US-10-369-493-27884

Query Match	16.3%;	Score 326.2;	DB 12;	Length 1143;
Best Local Similarity	61.2%;	Pred. No. 5.9e-72;		
Mismatches	545;	Conservative	0;	Mismatches 343;
Indels	3;	Gaps	1;	
QY	647	CCTATATAGGTTTGGATGACCTGCTTGCACATGATTTTATCTGAAATCTGGTAGAT	706	
DB	247	CAATTATAGGTTTGGAGCCCTGAGCTGTACAGCTACTACTACTACTAGTTCGAC	306	
QY	707	TGGTCTTCACACATGTTGGCTGTTGGTTGGTTAACTGTGTCTCTCTGAAAGCT	766	
DB	307	TGGGGAGCGCCACAGTCTCTGGGCTTGGTTTAGGATCCAGCGTTTACATCTGGAATGCG	366	
QY	767	TGTACACAGGTTAACTAAATTAATGATTTGGGGTGTGATGTTGTTGTTCTGTT	826	

Db 367 CAACAAAGTAGAGTCAACAAAGCTGTGCACGTTGGAG---GACGACACGGTGGCCACGCTA 423
Qy 827 GGTGGGCTCAACGCTGGTACTCACTCTGTGTTGGAACTAAACAAATGGTAAAGTTCAGATT 886
Db 424 TCTTGGATACAAAGGGGACACACCTTGTCTATTGGTACCCATTAAGGACTGTGTGCAGATC 483
Qy 887 TGGGATGACGACAGATGCAAGAGATGAAGATCAATGAGGGGCCATCGTTAGCTGTCGGG 946
Db 484 TGGGATGCTGAGAAAGCAAGAGCTTAAGAACGATGACTGCGCCATCTACTGCGAGGTTGGA 543
Qy 947 GCCTTGGCCCTGGAGTTCATCTCTTTTGTCTTCTGTGGACGGGATAGAAATATTTATCAA 1006
Db 544 GCATTGGCATGGAACACTCATATTCTCACTGAGGATCCCGGATCCGCTTAATATACCAT 603
Qy 1007 CGAGATATACGACACAGAGAGATTTGTTAGTAACTGTGACGACACAAATCAGAGGTT 1066
Db 604 CGGACCTCGAGGCGCAGATCAGTGGCTCAAAAGCTGTGTTGTCACAGCAGGAGTG 563
Qy 1067 TGTGGACTGAAGTGGTCAATATGATAACCGTGAAGTGGCACTCTGAGGAGAAATGACAA 1126
Db 664 TGTGGTCTCAATGGAATTGCGAGATGGCCAACTTGCAGTGGAGGAAACGACAAATAG 723
Qy 1127 TTGTTGTTGGAAATCAACACTCAACCCAGCTGTCTCAAGTACTGTGAGCACACAGCA 1186
Db 724 TTGATGCTTGGGATAGCTTTCCGACACACCTCTATGGAAGTATTCCGGTCAACAGCA 783
Qy 1187 GCTGTTAAAGCTATTGCAATGCTCTCTCATCTTCACTGCACTTCTTGGCATCTGAGGAGGA 1246
Db 784 GCCGTCAAGCGATAGCTGGTCAACACATCAAGCGGTCTCTGGCTTCCGAGGCGGT 843
Qy 1247 ACTGACAGATGATGTTTGGTTTGGAAATCAACACCAAACTCAACCTCACTGCTGTATG 1306
Db 844 ACTGCTGACAGACCATCATCTTTTCCAGCACACCGTTCCGCGCACGGTACTCAACGAGGTG 903
Qy 1307 GACACTGGAAGTCAAGGTTTGGAACTTGTCTGTGCTCCAAATGTCACAGAACTAGTAAGC 1366
Db 904 GACACGGGCTGAGTCTGTAACATTGCTGTGCTCCAGAACTCCAAAGAAATGCTGCTG 963
Qy 1367 ACACATGGGTACTCCGAGAACAGATTAATGTTTGGAGATACCCCACTATGTCAAAGCTG 1426
Db 964 ACGCATGGATATAGCCAGAACAGATAGTATTTTGGAAAGTACCCCTCCATGACGCAAGTC 1023
Qy 1427 GCGACTCTACCGGCTACTTATAGGTTCTCTATCTTGGCATCTCTCCAGATGGACAG 1486
Db 1024 GCCAGCTGACCGGTCACTATTCGGGTGCTGTACTTGGCATGACCCCGGACGCAAA 1083
Qy 1487 ACTATTGTAACTGGAGCTGGAGATCAAAACGCTTAGGTTCTGGAATGTTTC 1537
Db 1084 ACCGTGGTACCGGTGCGGAGATGAGACATTGGGGTTCTGGAATCTCTTC 1134

RESULT 2
US-10-032-585-6595
; Sequence 6595, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6595
; LENGTH: 1794
; TYPE: DNA
; ORGANISM: Candida albicans
US-10-032-585-6595

Query Match 15.1%; Score 303.8; DB 13; Length 1794;

Best Local Similarity 58.1%; Pred. No. 3.6e-66;
Matches 536; Conservative 0; Mismatches 387; Indels 0; Gaps 0;
Qy 614 CCGGTTAAGGCTCTAGGAAGGTTCTCGATCGGCTTATAAGGTTTGGATGACCTGCT 673
Db 805 CCGCAGAAGAAAGCCCTCGAAACAATATCAAGGTTCCATATAGAGTATTGGATGCAACAGAA 864
Qy 674 TTGCAAGATGATTTTATCTGAATCTGTAGATTGGTCTTTCACAAATGTGTGGCTGTT 733
Db 865 CTATCTGATGATTTTATTTGAATTTAGTTGATTTGGGTCACACAGGATGACTAGCCGTA 924
Qy 734 GGTTTGGTAACTGTGCTATCTCTGGAATGCTGTAGCAGCAAGGTAACATAATTTATGT 793
Db 925 GGAATGGGCGACAGTGTGTATTATTTATGGGATGGGCCCAACAATCAGTGGACAGATTTGT 984
Qy 794 GATTTGGGGTGTGATGATTTGTTTCTCTGTTGGGCTCAACGCTGGTACTCATCTT 853
Db 985 AATTTCACAAACAGGCAAAAGTCAAGTTTGAATGGATAGGTACTGGTACCCATTG 1044
Qy 854 GCTGTTGGAATCAACATGGTAAAGTTCAGATTTGGGATGCGCAAGATGCAAGATGCAAGAATA 913
Db 1045 GCAATAGGAACATCAAGGGTTTGGTGGAGATATGGGATGCAACTCGAATCAAGTGCATC 1104
Qy 914 AGATCAATGGAGGGCCATCGGTTACGTTGCGGGCCTTGGGCTCGAGTTCACTCTTTTG 973
Db 1105 AGGACATGACAGGACACTCTCTACGGTTAGTTTCATTAGCGTGAATGAACATATTA 1164
Qy 974 TCTTCTGTTGACCGGATTAAGAAATTTATCAACGAGATATACGACACAAGAAGATTTT 1033
Db 1165 AGCAGTGGGAGTTCGGATCGAACCATTTTAAATCGGATGTGAGAAATTTGAAGATCCTTT 1224
Qy 1034 GTTAGTAACTGTGAGGACACAATCAAGCTTGTGGACTGGAAGTGCATATATGATAAC 1093
Db 1225 GTTAAATAATTTGATPAATCATAGCAAGAGTGTGGACTCAAGTGAATGTGGAAGAA 1284
Qy 1094 CGTGAAGTTGGCATCTGGAGGAATGACAAACAAATTTGTTTGTGGAATCAACCTCAACC 1153
Db 1285 AACAACTAGCTAGTGTGGTGGCAACGATAATAATTTGTTGATGGGATGGGCTCAATCCT 1344
Qy 1154 GAGCTGTCTCAAGTACTGTGACGACACAGCAGCTGTTAAAGCTATTGCAATGCTCTCT 1213
Db 1345 AAACGTTTGCATCAATTTACAGACCACTGCTGCCGTTAAAGCTATTGCTTGGTCCGG 1404
Qy 1214 CATCTTCATGAGCTTTCTGCACTCTGGAGGAGAACTGCAAGATAGATGTATTGTTTGG 1273
Db 1405 CATCAACGGGGATTTTGGCATCGGTCGCGAACAGCCGATAGACTATAAAGACATGG 1464
Qy 1274 AATACAAACCAAACTCAACCTTAGCTGTATGGACACTGGAAGTCAAGTTTGCATCTT 1333
Db 1465 AATACGTTGACTGGGAATTTTAGTACATGATGTCAATCTGCGCTCAAGTGTGTAATTA 1524
Qy 1334 GTCTGTCTCAAAAATGTCAACGAATAGTAAGCACACATGGGTACTCCAGAACCAAGATT 1393
Db 1525 ATCTGTCTAAAAAATTCACGAAATGGTTTCAACCCATGGTTATTCAGAAATCAAAAT 1584
Qy 1394 ATTGTTGGAGATACCCCACTATGTCAAGCTGGGACCTTTACCGGCCATCTTATAGG 1453
Db 1585 ATAGTGTGAAGTATCCATCAATGCAACAAATTTGCAAAATGACAGGTCATACATAGA 1644
Qy 1454 GTTCTCTATCTTGGCATCTCTCCAGATGACAGACTATTGTAACTGGAGCTGGAGATGAA 1513
Db 1645 GTTTTATTTTGTGTTGTTCTCTGATGTGAAACATCGTTACTTGTGCTGCTGATGAG 1704
Qy 1514 ACGTTAGTCTTGGAAATGTTTT 1536
Db 1705 ACCTTACGATTTTGAATGTATT 1727

RESULT 3
US-10-369-493-26005
; Sequence 26005, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:

APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 26005
LENGTH: 1312
TYPE: DNA
ORGANISM: Schizosaccharomyces pombe
US-10-369-493-26005

Query Match 13.8%; Score 277; DB 12; Length 1312;
Best Local Similarity 56.7%; Pred. No. 1.9e-59;
Matches 511; Conservative 0; Mismatches 390; Indels 0; Gaps 0;

QY 636 TTCTCGATCGCTTATAAGGTTTGGATGCACCTGCTTGCAGATGATTTTATCTCA 695
DB 372 TTCAAAACTCCATATAAATCTTGGATGCACCTTACTTAAAAATGACTTTTATTA 431
QY 696 ATCTGGTAGATTGGTCTTACACAATGTGTGGCTGTGGTAACTGTGTCTATC 755
DB 432 ACTTGTGACTGGGCGCAATCTAATGTCTTGGCAGTCGACTAGCAAGCAGCATTTATC 491
QY 756 TCTGGAATGCTTGTAGCAGCAAGGTAACTAAATATGTGATTTGGGGTGTGATTTGTG 815
DB 492 TCTGTCAGCAGCAAGTGGTAAAGTGGTTCAATTACATGATTCGGAGCAACAAATCATG 551
QY 816 TTTGTTCTTGTGGTGGGCTCAACGTGTACTCTCTTGTCTTGGTAACTAACTAGTGA 875
DB 552 TCACAAAGCGTCTTATGGCAGGTAAAGNACTCAATAGCTGTAGGACAGATTTCAGAG 611
QY 876 AAGTTTCAGATTTGGGATCAGCAAGATGCAAGAGATAAGATCAATGAGGGCCATTCGGT 935
DB 612 TAAATTTACATCTGGGATATCGAGTCTACGAAATCAGTCAGATCTTTAAAGGTCACTCCG 671
QY 936 TACGTGTGGGGCCCTTGGCTGGAGTTTCATCTCTTTTGTCTTCTGATTAACCGTGGAGGAA 1115
DB 792 AACAGAAATATGCGATTGCAATGGATAGAGTCTAGCAGCTTGTCTCAGGAGGGA 851
QY 1116 ATGCAACAAATTTGTTTGGATCAACACTCAACCCAGCTGTCTCAAGTACTGTG 1175
DB 852 ATGCAATAATTTATTCGTATGGGATTAACCGCTCTTCACTCTCTTACATAAGTTTGAAG 911
QY 1176 AGCACAGCAGCTTTAAAGCTATTTCATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1235
DB 912 AGCATACAGACCGGTCAAGCAATTTGGATGGAGCCCTCAACAGGGGATCTCGCA 971
QY 1236 CTGGAGGAGGAACCTGCAGATAGATGTATTCGTTTGGGAATACAAACCAAACTCAACC 1295
DB 972 CGCGCGGGGTACCATCGACAGATGCTTAACGATTCATAATATCTTTGACTGGAAGGCTAC 1031
QY 1296 TTAGCTGTAGGACCTGGAGTTCAGTTTGGATCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 1355
DB 1032 AAAATAAGTTGATCTGGTCTCAAGATGTACATGGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 1091
QY 1356 AACTAGTAAGCACATCGGTGCTCTCCAGAACCCAGATTAATGTTTGGAGATACCCACCTA 1415

DB 1092 AAATGTTACTACTATGTTTGGAAAAAACCAAGATATCCTTATGGAATATCCTTCAT 1151
QY 1416 TGTCAAAGCTGGGACTCTTTACCGGCCATACTTATAGGTTCTCTATCTTCCCATCTCTC 1475
DB 1152 TGAATAATATCGCAATCTTACTTGCCTATCAAAACCGTGTCTCTTCTATGAGTC 1211
QY 1476 CAGATGACAGACTTATGTAACCTGGAGCTGAGATGAACGCTTAGGTTCTTGGAACTTT 1535
DB 1212 CCGATGCCAGTCTATAGTTACGGGAGCTGGGACGAAACGCTAAGATTTTGGAGCTTT 1271
QY 1536 T 1536
DB 1272 T 1272

RESULT 4
US-10-369-493-26011
Sequence 26011, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 26011
LENGTH: 1312
TYPE: DNA
ORGANISM: Schizosaccharomyces pombe
US-10-369-493-26011

Query Match 13.8%; Score 277; DB 12; Length 1312;
Best Local Similarity 56.7%; Pred. No. 1.9e-59;
Matches 511; Conservative 0; Mismatches 390; Indels 0; Gaps 0;

QY 636 TTCTCGATCGCTTATAAGGTTTGGATGCACCTGCTTGCAGATGATTTTATCTCA 695
DB 372 TTCAAAACTCCATATAAATCTTGGATGCACCTTACTTAAAAATGACTTTTATTA 431
QY 696 ATCTGGTAGATTGGTCTTACACAATGTGTGGCTGTGGTAACTGTGTCTATC 755
DB 432 ACTTGTGACTGGGCGCAATCTAATGTCTTGGCAGTCGACTAGCAAGCAGCATTTATC 491
QY 756 TCTGGAATGCTTGTAGCAGCAAGGTAACTAAATATGTGATTTGGGGTGTGATTTGTG 815
DB 492 TCTGTCAGCAGCAAGTGGTAAAGTGGTTCAATTACATGATTCGGAGCAACAAATCATG 551
QY 816 TTTGTTCTTGTGGTGGGCTCAACGTGTACTCTCTTGTCTTGGTAACTAACTAGTGA 875
DB 552 TCACAAAGCGTCTTATGGCAGGTAAAGNACTCAATAGCTGTAGGACAGATTTCAGAG 611
QY 876 AAGTTTCAGATTTGGGATCAGCAAGATGCAAGAGATAAGATCAATGAGGGCCATTCGGT 935
DB 612 TAAATTTACATCTGGGATATCGAGTCTACGAAATCAGTCAGATCTTTAAAGGTCACTCCG 671
QY 936 TACGTGTGGGGCCCTTGGCTGGAGTTTCATCTCTTTTGTCTTCTGATTAACCGTGGAGGAA 995
DB 672 AAAGGTGGCAGCATTTAGCATGGAAACGATATACCTCAAGTGGAGGCAAGACGAG 731
QY 996 ATATTTATCAACGAGATATAGCCACACAAGAAAGATTTTGTAGTAACTGTCAAGACACA 1055
DB 732 TAAATTTACATCTGGGATTAACCGCTCTTCAAGTGTACATGGCTGTCTCTCTCTCTCTCTCT 791
QY 1056 AATCAGAGTTTGGGATCTGAGTGGTCTATGATTAACCGTGGAGTTCGATCTCGAGGAA 1115

Db 792 AACAGAAATATCGGATTCGATGATGATAGAGCTTAGACAGCTTGTCTTCAGAGGGA 851
Qy 1116 ATGACACAAATCTTTGTTGGATCAACACTCAACCCAGCTGTCTCAAGTACTGTG 1175
Db 852 ATGACAAATATTTATTCGTATGGGATACCGCTCTTCAAGTCTTTACATAGTTGAAG 911
Qy 1176 AGCAGACAGAGCTGTTAAAGCTATTGATGCTCTCTCTCATCTTTCATGGAATCTTCTG 1235
Db 912 AGCATACAGAGCGGTCAAGCAATTTGGATGGAGCCCTCATCAACGAGGGATCTTCGCA 971
Qy 1236 CTGAGGAGGAACTGCAGATAGATGATTCGTTTGGATGATACCAACCAAACTCACACC 1295
Db 972 GCGGCGGGGTACCATGACAGATGCTTAAGCATATATATCTTGTGGAAGGCTAC 1031
Qy 1296 TTAGCTGTATGAGACTGGAAGTCAAGTTTGAATCTTGTCTGTCGTCGTCGTCGTCGTC 1355
Db 1032 AAAATAAGTTGATGACTCGCTCAAGTATGTAACATGATGTAACATGTAACATGTAAC 1091
Qy 1356 AACTAGTAGACACATGCTGCTCAAGTATGTAACATGTAACATGTAACATGTAACATGTA 1415
Db 1092 AAAATGTTACTACTGCTCAAGTATGTAACATGTAACATGTAACATGTAACATGTAAC 1151
Qy 1416 TGTCAAGCTGCGGACTCTTACCGGCTCATCTATAGGTTCTCTATCTTGTGCTCTCTC 1475
Db 1152 TGAAATATGCTCAATCTTACTGCTCAACATGTAACATGTAACATGTAACATGTAAC 1211
Qy 1476 CAGATGAGACATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1535
Db 1212 CCGATGGCAGCTATAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1271
Qy 1536 T 1536
Db 1272 T 1272

RESULT 5

US-10-369-493-26013
; Sequence 26013, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 26013
; LENGTH: 1312
; TYPE: DNA
; ORGANISM: Schizosaccharomyces pombe
US-10-369-493-26013

Query Match 13.8%; Score 277; DB 12; Length 1312;
Best Local Similarity 56.7%; Pred. No. 1.9e-59;
Matches 511; Conservative 0; Mismatches 390; Indels 0; Gaps 0;
Qy 636 TTCCTCGATCGCTTAAAGTTTGGATGACCTGCTTGGCAAGATGATTTTATCTGA 695
Db 372 TTCAAAAATCCATATAAATCTTGGATGACCTTACTTAAAAAATGACTTTTATTA 431
Qy 696 ATCTGGTATGCTTTCACAAATGTTGCTGTTGTTGGTAACTGCTATC 755
Db 432 ACTTGCTGCTGGGCAATCTTAATGCTTGGAGTGGAGTACGACCATTTATC 491
Qy 756 TCTGGAATGCTTGTACAGCAAGTAATATGATTTGGGGTGTGATTTG 815

Db 492 TCTGCTAGCAGCAAGTGGTAAAGTGTTCATTAATGATTTCCGAGCAACAATCATG 551
Qy 816 TTTGTTCTGTTGGTCTCAACCTGCTACTCATCTTCTGTTGGAACTCAACATGTA 875
Db 552 TCACAGCGTCTTATGAGCAGGTAAGGAACCTCAATTAGCTGTAGGACAGATTCCAGG 611
Qy 876 AAGTTCAATTTGGGATGACCAAGATGCAAGAGATGATCAATGAGGCGCATCGT 935
Db 612 TAATTTACATCTGGGATATCGAGTCTAGCAATCATGTCAGATCTTTAAAGTCTACTCG 671
Qy 936 TAGCTGTCTGGGCTTGGCTGAGTTCATCTCTTTTCTTCTGTTGGAGCGGTAAGA 995
Db 672 AAGGGTGGCAGCAATTAGCATGGAACGATAATACCTCACAAAGTGGAGCAAGAGCAAG 731
Qy 996 ATATTTATCAAGATATACCAACAGAGATTTTGTAGTAACTGTCCAGACACA 1055
Db 732 TAATTTTACATCATGACCTTAAGAGCAGGATGCTGTGCCGAAATGATGAAGTTTCA 791
Qy 1056 AATCAGAGGTTTGTGAGCTGAAAGTGTCTATATGATTAACCGTGGAGTTGGCATCTG 1115
Db 792 AACAGAAATATGCGGATTTGCAATGGATAGAGTCTAGGACAGCTTGTCTTCAGGAGGA 851
Qy 1116 ATGACAAATTTTGTGGAATCAACACTCAACCCAGCTCTCTCAAGTACTGTG 1175
Db 852 ATGACAAATTTTGTGGAATCAACACTCAACCCAGCTCTCTCAAGTACTGTG 911
Qy 1176 AGCAGACAGCTGTTAAAGCTATTGATGCTCTCTCATCTTCAATGACTTCTTGCAT 1235
Db 912 AGCATACAGCGGTCAAGCAATTTGGATGAGCCCTCATCAACGAGGATCTTCGCA 971
Qy 1236 CTGAGGAGAACTGAGATAGATGATGATGATGATGATGATGATGATGATGATGATG 1295
Db 972 GCGGCGGGGTACCATCGACAGATGTTAAACGATTCATAATATCTTGAAGGCTAC 1031
Qy 1296 TTAGCTGTATGAGACTGGAAGTTCAGTTTGAATCTTGTCTGCTGCTCAAAATGTCAAG 1355
Db 1032 AAAATAAGTTGATGCTGCTCAAGTATGTAACATGATGCTGCTGCTGCTGCTGCTGCT 1091
Qy 1356 AACTAGTAGACACATGCTGCTCAAGTATGTAACATGATGATGATGATGATGATGATG 1415
Db 1092 AAATTTGTTACTACTGCTTTCGCAAAATGCAAGTATCTTATGGAATATCTTCTCAT 1151
Qy 1416 TGTCAAGCTGCGGCTTACCGGCTCATCTATAGGTTCTCTATCTTGTGCTCTCTC 1475
Db 1152 TGAAATATGCGCAATCTTATGCTCCATCAACCGTCTCTCTTCTTCTTCTGATC 1211
Qy 1476 CAGATGAGACAGATTTGTAATGAGCTGGAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1535
Db 1212 CCGATGGCAGTCTATAGTATGAGGAGCTGGGAGCGGAGCGGAGCGGAGCGGAGCT 1271
Qy 1536 T 1536
Db 1272 T 1272

RESULT 6

US-10-369-493-26526
; Sequence 26526, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 36526
; LENGTH: 1116
; TYPE: DNA
; ORGANISM: Aspergillus nidulans
; US-10-369-493-36526

Query Match 13.2%; Score 264.6; DB 12; Length 1116;

Best Local Similarity 57.7%; Pred. No. 2.4e-56;
Matches 542; Conservative 0; Mismatches 349; Indels 48; Gaps 2;

```
QY 647 CTTTAAAGTTTGGATGACCTGCTTTGCAAGATGATTTTATCTGAATCTGGTAGAT 706
Db 178 CCCTACAAGTTTCTCGATGCCAGACTTTCAGGACGATTTTACCTGAATTTGGTTGAT 237
QY 707 TGGTCTTACACAAATGTTGGCTGTTGGTTTGGGTAACCTGCTCTATCTCTGGAATGCT 766
Db 238 TGGGGAGTAGTAGTTCTAGCGTTGGCTTAGGAACTCGGTGTACATGTGGAACTCA 297
QY 767 TGTAGCAGCAAGTAATAATATATGATTTTGGGGTTG----- 806
Db 298 CAAACTGGGAGGTTTACGAAATTTGTGTAGCTTTAAGGATGACACAGTCAGCGCTCACT 357
QY 807 -----ATGATTTGTTTGTCTTGTGTTGTTGTTGTTGGCTCAAC 839
Db 358 TGGATCAAAAGGTAGCCAGCTCTTATGAAATGTTCTGGAGCTTATGTTGATAACATC 417
QY 840 -GTGCTACTCATCTGCTTTGGAATCAAAATGTTAAAGTTTCAGATTTGGGATGCAGCA 898
Db 418 TAGGGTACGCACCTTTCAAATGTTGACGGGAAAGGTATGTTGCAAAATATGGGATGCAGAG 477
QY 899 AGATGCAAGAAGATAAGATCAATGAGGAGCCATCGGTAGCTGTCGGGCTTGGCCTGG 958
Db 478 CGCTGTCCGCCCTACGGACCATGTGGGCACCAATCGCTAGGGCGGTTGGCTTGG 537
QY 959 AGTTCATCTCTTTGTTCTTCTGTTGGACGGGATAAGAAATTTATCAACAGATATACGC 1018
Db 538 AACGATCATATCTGACATCGGCTCTCGGATCGGCATATTTTTCATCGTACGTCGG 597
QY 1019 ACACAAGAAGTTTGTAGTAATCTCTCAGGACACAAATCAGAGTTTGTGAGCTGAAG 1078
Db 598 TCTCCTGACCAGTATCTTCTGCTGCACTTCTGCTCATAGCAGGAAGTGTGGGCTCAGG 657
QY 1079 TGGTCATATGATAACCGTGAAGTTGGCATCTGGAGGAAATGACAACTAAATGTTGTTGG 1138
Db 658 TGGAAACAGGAGATGTTCACTGGCATCAGGGGCAACGACATAAGCTCTCTGTTGG 717
QY 1139 AATCAACACTCAACCCAGCTGCTCCTCAAGTATCTGTGAGCACACAGCAGCTGTTAAAGCT 1198
Db 718 GACAAATTGAACGAGACCCCTCTTTATCGCTTCTCAGACCATACCGCGCCGCTGAAGCT 777
QY 1199 ATTGCATGGTCTCTCATCTTTCATGGACTTCTTGCATCTGAGGAGAACTGCAGATAGA 1258
Db 778 ATCATGTTTCAACCCATCAACACCATTTACTCGCATCCGGGGAGGTACTGCTGATCGA 837
QY 1259 TGTATTCGTTTTTGGAAATACAAACCAAACTCACACCTTAGCTGTATGGACACTGGAAGT 1318
Db 838 ACGATTAAATTTTGGAAACAGCCGACGGTTCTCTTATCAAGGAGTTGATACCGGAGC 897
QY 1319 CAGGTTTCATCTTGTCTGTCGCAAAATGTCACGAACTAGTAAAGCACACATGGGTAC 1378
Db 898 CAAGTCTGTAACTTGGCATGCTGCAAGATTTCTGATGAATATTCAGTACGATGGCTAT 957
QY 1379 TCCAGAACCAAGATTAATTTGTTGGAGATACCCCACTATGTCAAAGCTGGCGACTCTTACC 1438
Db 958 AGTCAGAACCAAAATTTGTCATCTGGAAGTATCTCTGATGGAGCAGATTTGTGCTTACG 1017
QY 1439 GGCCATCTTATAGGGTTCTTATCTTGGCACTCTCCAGATGACAGACTATTTGTAATC 1498
Db 1018 GGCCATCTTATAGGGTTCTTATCTTGGCACTCTCCAGATGACAGACTATTTGTAATC 1557
QY 1499 GGAGCTGAGATGAACCGCTTAGGTTCTGGAATGTTTTC 1537
Db 1078 GGTCCCGCGCAGACCCCTACGGTTCTGGAAGATATTC 1116
```

RESULT 7

US-10-128-714-1038
; Sequence 1038, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:

; APPLICANT: Jiang, Bo

; APPLICANT: Hu, Wenqi

; APPLICANT: Tishkoff, Daniel

; APPLICANT: Zamudio, Carlos

; APPLICANT: Broshkin, Alexey M

; APPLICANT: Lemieux, Sebastien M

; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and

; FILE REFERENCE: 10182-018-999

; CURRENT APPLICATION NUMBER: US/10/128,714

; CURRENT FILING DATE: 2002-04-23

; PRIOR APPLICATION NUMBER: US 60/285,697

; PRIOR FILING DATE: 2001-04-23

; PRIOR APPLICATION NUMBER: US 60/287,066

; PRIOR FILING DATE: 2001-04-27

; PRIOR APPLICATION NUMBER: US 60/295,890

; PRIOR FILING DATE: 2001-06-05

; PRIOR APPLICATION NUMBER: US 60/303,899

; PRIOR FILING DATE: 2001-07-09

; PRIOR APPLICATION NUMBER: US 60/316,362

; PRIOR FILING DATE: 2001-08-31

; NUMBER OF SEQ ID NOS: 8603

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1038

; LENGTH: 1398

; TYPE: DNA

; ORGANISM: Aspergillus fumigatus

US-10-128-714-1038

Query Match 13.2%; Score 263.8; DB 15; Length 1398;

Best Local Similarity 56.3%; Pred. No. 4.3e-56;

Matches 557; Conservative 0; Mismatches 372; Indels 60; Gaps 1;

```
QY 609 ATAGTCCGGTTAAGCTCTCTAGGAGGTTCCCTCATCGCCTTATAAGGTTTGGATGCAC 668
Db 341 AGACTCTCCGCAAGCAGCTCGCTAATAAGTACCCCTATAAGGTTTCCGACGCAC 400
QY 669 CTGCTTTGCAAGATGATTTTATCTGAACTCTGTAAGTTGCTTTCACAAATGTGTTGG 728
Db 401 CGGATTTGCAAGATGATTTTCTACTTGAATCTGTTGCACTGGGCGAGTAGTAATGTCTAG 460
QY 729 CTGTTGGTTTGGGTAACCTGTTCTATCTCTGAAATGTTGTAGCAGCAAGTAATAAT 788
Db 461 GTGTTGGGCTTAGGCAATTCAGTGTACATGTGGAATTCGAATACCGGACGGGTGACGAAAC 520
QY 789 TATGTGATTTGGGGTTGATGATTG----- 813
Db 521 TTTGCAACTAAGATGATGATCTGTCACAAGTTGTAGCTGGATTCAGAGGGTAATGAGAT 580
QY 814 -----TGTTGTTCTGTTGGTGGGCTCAACGTGGTACTC 848
Db 581 TTCTCTATATGTTGCAATGATTATATGTTGTTTCACTGACTGACATTTGGACAGGGCACAC 640
QY 849 ATCTTCTGCTTGGAACTAACAATGTAAGTTTCAGATTTGGGATTCAGCAAGATCAAGA 908
Db 641 ATCTTTTCAATAGAACTGGCAAGGTCCTGTACAGATATGGATGCGAGCATTTGTCGCC 700
QY 909 AGATAAGATCAATGGAGGGCCATCGGTTACCGTCTCGGGCCCTTGGCTGGAGTTTCACTC 968
Db 701 GTCTTCGGACAATGATTGGGCATACCAATCGTGTGGGGCTCTTCTTGGAACTGATCATA 760
QY 969 TTTTGTCTCTGTTGGAAGGATAGATATTTATCAACGAGATATAGCACAAGAGAG 1028
Db 761 TCTTTACGTCAGGTTTCTCGGATCGACTGATCTTCCACCGTGTGTTCTGTTCCCGCATC 820
QY 1029 ATTGTTGTTAGTAAATGTTTCAGGACACAAATCAGAGGTTTGTGGACTGAATGTTGATATG 1088
```


821	AGTATTTACGTGCAGTGTCCGGCCATATGACGAGAAAGTTTGGCGACTCCGGTGGAAACACCG	980
1089	ATAACCGTGAGTTGGGCATCTGGAGGAAATGACAACAAATGTTTGTTTGGAAATCAACACT	1148
881	AAGATGGCCAACTAGCTTTCAGCGCGGAAATGACAACTCATGCGTGTGGGCAAGCTGA	940
1149	CAACCCGAGCTGTCCTCAAGTACTGTGTAGCACACAGCAGCTGTATAAAGTATTTCATGCT	1208
941	ATGAGACACCGCTTTATCGCTTCTCGGACCATACCGCGCTGAAGGCCATCGCCTGGT	1000
1209	CTCCTCATCTTATGCACTTCTGCACTCTGGAGGAGGAACTGCAGATAGATGTATTTCGTT	1268
1001	CACCCCATCAACACCATCTCCTCGCTCAGTGGAGGACGCGGCGACCGACCATCAAAAT	1060
1269	TTTGGAAATCAACCAAACTCAACACTTACGCTGTATGGACACTGGAAGTCAGGTTTGCA	1328
1061	TCGTGGAACACTGCTACTGTTTCACTGATCAAGGAAGTTGATACGGGTAGCCAGGTGTGA	1120
1329	ATCTTGTCTGGTCCAAAATGTCACGAACTAGTAAAGCAACATGGGTACTCCAGAAC	1388
1121	ACCTGGCATGGTCCGAGAACTCGGAGGAAATCATCAGTACACACGGCTCAGTTCGAAAC	1180
1389	AGATTATTGTTTGGAGATACCCCACTATGTCAAAGCTGGCGACTCTTACCGGCCATACCT	1448
1181	AGATTGTAATCTGAAAATACCCGGCATGGAGCAGATTTGATCGCTACCGGCACACATT	1240
1449	ATAGGGTCTCTATCTTGCCATCTCTCCGATCGGACAGACTATTGTAACTGGAGCTGGAG	1508
1241	TCGGTGTCTCTACCTTTGCGATGAGCCCGCATGGACAGACCGTCTGTACCGGAGCTGGTG	1300
1509	ATGAAACGCTTAGGTTCTCGAATGTTTTTC	1537
1301	ACGAAACGTTGAGATCTCGAAGATTTTC	1329

RESULT 8

US-10-128-714-2038
 ; Sequence 2038, Application US/10128714
 ; Publication No. US20030119013A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jiang, Bo
 ; APPLICANT: Hu, Wenqi
 ; APPLICANT: Tishkoff, Daniel
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Eroshkin, Alexey M
 ; APPLICANT: Lemieux, Sebastien M
 ; TITLE OF INVENTION: Identification of Essential Genes in *Aspergillus fumigatus* and
 ; TITLE OF INVENTION: Methods of Use
 ; FILE REFERENCE: 10182-018-999
 ; CURRENT APPLICATION NUMBER: US/10/128,714
 ; CURRENT FILING DATE: 2002-04-23
 ; PRIOR APPLICATION NUMBER: US 60/285,697
 ; PRIOR FILING DATE: 2001-04-23
 ; PRIOR APPLICATION NUMBER: US 60/287,066
 ; PRIOR FILING DATE: 2001-04-27
 ; PRIOR APPLICATION NUMBER: US 60/295,890
 ; PRIOR FILING DATE: 2001-06-05
 ; PRIOR APPLICATION NUMBER: US 60/303,899
 ; PRIOR FILING DATE: 2001-07-09
 ; PRIOR APPLICATION NUMBER: US 60/316,362
 ; PRIOR FILING DATE: 2001-08-31
 ; NUMBER OF SEQ ID NOS: 8603
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 2038
 ; LENGTH: 1398
 ; TYPE: DNA
 ; ORGANISM: *Aspergillus fumigatus*
 ; US-10-128-714-2038

Query Match 13.2%; Score 263.8; DB 15; Length 1398; GC 40.2%;
Best Local Similarity 56.3%; Pred. No. 4.3e-56;
Matches 557; Conservative 0; Mismatches 372; Indels 60; Gaps 1;

RESULT 9
US-10-128-714-7038
; Sequence 7038, Application US/10128714

Publication No. US20030119013A1
GENERAL INFORMATION:
APPLICANT: Jiang, Bo
APPLICANT: Hu, Wendi
APPLICANT: Tishkoff, Daniel
APPLICANT: Zamudio, Carlos
APPLICANT: Eroshekin, Alexey M
APPLICANT: Lemieux, Sebastien M
TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
TITLE OF INVENTION: Methods of Use
FILE REFERENCE: 10182-018-999
CURRENT APPLICATION NUMBER: US/10/128,714
CURRENT FILING DATE: 2002-04-23
PRIOR APPLICATION NUMBER: US 60/285,697
PRIOR FILING DATE: 2001-04-23
PRIOR APPLICATION NUMBER: US 60/287,066
PRIOR FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/295,890
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: US 60/303,899
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: US 60/316,362
PRIOR FILING DATE: 2001-08-31
NUMBER OF SEQ ID NOS: 8603
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7038
TYPE: DNA
ORGANISM: Aspergillus fumigatus
US-10-128-714-7038

Query Match 13.2%; Score 263.8; DB 15; Length 1875;
Best Local Similarity 56.3%; Pred. No. 5.1e-56;
Matches 557; Conservative 0; Mismatches 372; Indels 60; Gaps 1;
QY 609 ATAGTCGGTTAGGCTCTAGGAAGGTCCTCGATCGGCTTATTAAGGTTTGGATGCAC 668
DB 818 AGACTCTCGCAAGCAGCTCGTACGTTAATAAGTACCCATATAGGTTCTCGACGCAC 877
QY 669 CTGCTTGGCAAGATGATTTTATCTGAATCTGTAGATGTTCTTACCAATGTTGG 728
DB 878 CCATTTGCAAGATGATTTTATCTGAATCTGTAGATGTTCTTACCAATGTTGG 937
QY 729 CTGCTTGGTTAGGTAATGTTCTTCTGGAATGTTCTGATGACGACGAAGTAATTAAT 788
DB 938 GTGTTGGCTTAGGCAATTCAGTCTGATGTTGGAATTCGAATACCGACGGGTGACGAAC 997
QY 789 TATGTAATTTGGGGTGTGATGTTG-----TGTTTGTCTGTTGGTGGGCTCAACGTGTAATC 813
DB 998 TTTGCGAATTAAGATGATGATCTGTCACAGTGTAGCTGATTCAGAGGTAATGAT 1057
QY 814 -----TGTTTGTCTGTTGGTGGGCTCAACGTGTAATC 848
DB 1058 TTCTCTATATGTTGGCAATGATTTATGTTGTTTCAATGACTGACATTCGACAGGCGACAC 1117
QY 849 ATCTTGTGTTGAACTAACAAGTTAAAGTTGAGTTGGATGCGACGACGATGCAAGA 908
DB 1118 ATCTTTCAATAGGAATGCGAAGGTTCTGACAGATATGGATGCGACGATTCGCGC 1177
QY 909 AGATAAGATCAATGAGGGGCGCATCGTTACGTTGCTCGGGGCTTGGCTGAGTTCAATCTC 968
DB 1178 GTCTTCGCAATGATTTGGGCATACCAATCGTGTGGGGGCTCTTGTTCGAACGATCA 1237
QY 969 TTTTGTCTTCTGGTGGCGGATAGAAATTTATCAAGAGATATACGACACAAGAG 1028
DB 1238 TCCTTAGCTCAGGTTCTCGGGATCGACTGATCTTCCACCGTGTATGTTCTCCCGCATC 1297
QY 1029 ATTTTGTAGTAACTCTGAGGACACAAATCAGAGGTTTGTGACGTGAAGTGTCTATAG 1088
DB 1298 AGTATTTACGTCAGTCTCGGGCCATAGCAGGAAGTTTGGGACTCCGGTGAACCG 1357
QY 1089 ATACCGTGTAGTGGCAATCTGAGGGAATGACAAATTTGTTTGTGTAATCAATCACT 1148

DB 1358 AAGATGGCCAACTAGCTTTCAGGGCGGAATGACAACTCATGTTGGGACAAGCTGA 1417
QY 1149 CAACCCAGCTGTCTCTCAAGTACTGTGAGCACACAGCAGCTGTAAAGCTATTGCTATG 1208
DB 1418 ATGAGACACCGCTTTATTCGCTTCTCGSACCATACCGCAGCCGTGAAGGCCATGCGCTG 1477
QY 1209 CTCTCATCTTCATGGACTTCTTGCCATCTGAGGAGGAACCTGCAGATAGATGATGTTG 1268
DB 1478 CACCCCATCAACCATCTCTCGCTCAGGTGGAGGACCGGGGACCGGACCATCAAT 1537
QY 1269 TTTGGAATACAAACCAACAACTCACACCTTAGCTGTATGAGACACTGGAAGTCAGGTTT 1328
DB 1538 TCTGGAACACTGCTACTGTTCTACTGATCAAGGAAGTTGATACGGGTAGCCAGGTGTA 1597
QY 1329 ATCTTGTCTGTCGCAAAATGTCACCAACTAGTAAACACACATGCTGCTATCCAGAAC 1388
DB 1598 ACCTGGCATGCTCGAAGAACTCGGACGAATATCATAGTACACAGGCTACAGTCAGAAC 1657
QY 1389 AGATTATTGTTGGAGATACCCCACTATGTCAAAGCTGGGCACTCTTACCGGGCCATACT 1448
DB 1658 AGATTGTAATCTGGAATACCGCGCATGGAGCAGATTGTATCGTACCGGCCACACTT 1717
QY 1449 ATAGGGTCTCTATCTTGGCATCTCTCCAGATGAGACAGACTATTGTAATCTGAGCTGG 1508
DB 1718 TCCGTGTTCTCTACCTTGGCATGAGCCCGATGACAGCCGCTGTTTACCGGAGCTG 1777
QY 1509 ATGAACGCTTAGCTTCTGGAATGTTTTC 1537
DB 1778 ACGAAGCTTGAGATTCTGGAAGATTTC 1806

RESULT 10
US-10-128-714-6038
Sequence 6038, Application US/10128714
Publication No. US20030119013A1
GENERAL INFORMATION:
APPLICANT: Jiang, Bo
APPLICANT: Hu, Wendi
APPLICANT: Tishkoff, Daniel
APPLICANT: Zamudio, Carlos
APPLICANT: Eroshekin, Alexey M
APPLICANT: Lemieux, Sebastien M
TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
TITLE OF INVENTION: Methods of Use
FILE REFERENCE: 10182-018-999
CURRENT APPLICATION NUMBER: US/10/128,714
CURRENT FILING DATE: 2002-04-23
PRIOR APPLICATION NUMBER: US 60/285,697
PRIOR FILING DATE: 2001-04-23
PRIOR APPLICATION NUMBER: US 60/287,066
PRIOR FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/295,890
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: US 60/303,899
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: US 60/316,362
PRIOR FILING DATE: 2001-08-31
NUMBER OF SEQ ID NOS: 8603
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6038
LENGTH: 1993
TYPE: DNA
ORGANISM: Aspergillus fumigatus
US-10-128-714-6038

Query Match 13.2%; Score 263.8; DB 15; Length 1993;
Best Local Similarity 56.3%; Pred. No. 5.3e-56;
Matches 557; Conservative 0; Mismatches 372; Indels 60; Gaps 1;
QY 609 ATAGTCGGTTAAGGCTCTAGGAAGGTTCTCGATCGGCTTATTAAGGTTTGGATGCAC 668
DB 936 AGACTCTCGCAAGCAGCTCGTACGTTAATAAGTACCCATATAGGTTCTCGACGCAC 995

669 CTGCTTTGCAAGATGATTTTATCTGAATCTGTGATAGTGTCTTTCACAAATGTGTGG 728
Db CCGATTTGCAAGATGATTTTATCTGAATCTGTGATAGTGTCTTTCACAAATGTGTGG 1055
729 CTGCTTTGCAAGATGATTTTATCTGAATCTGTGATAGTGTCTTTCACAAATGTGTGG 788
Db GTGCTTTGCAAGATGATTTTATCTGAATCTGTGATAGTGTCTTTCACAAATGTGTGG 1115
789 TATGTGATTTGGGGTGTGATGTTG----- 813
1116 TTTGCGAATAAGAGATGATGATGTCACAAAGTGTGATGTTGATGTTGATGATGATGAT 1175
814 -----TGTTTCTCTGTTGTTGGCTCAACGGTGTACTC 848
1176 TTCTCTATATGTTGGCAAGATGATGTTGTTGTTTCACTGATGATGATGATGATGATGAT 1235
849 ATCTTGTCTGTTGGCAATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 908
1236 ATCTTGTCTGTTGGCAATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1295
909 AGATTAAGATCAATGAGAGGCGATCGTTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 968
1296 GTCTTTCGACATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1355
969 TTTTGTCTTCTGTTGGCAATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1028
1356 TCTTACGTCAGTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1415
1029 ATTTTGTGTTAGTAACTGTCAGCAACAATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1088
1416 AGTATTTACGTCAGTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1475
1089 ATAAACGTTGAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1148
1476 AAGATGGCACTAGTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1535
1149 CAACCCGCTGTTCTCAAGTACTGTGAGCAACAAGTGTGTTGTTGTTGTTGTTGTTGTTGTT 1208
1536 ATGAGACACCGCTTTTATGCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1595
1209 CTCTCTCATCTTCAAGTACTGTGAGCAACAAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1268
1596 CAACCCGCTGTTCTCAAGTACTGTGAGCAACAAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1555
1269 TTTGGAATCAACCACTCAACGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1328
1656 TCTGGAACACTGTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1715
1329 ATCTTGTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1388
1716 ACCTGGCATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1775
1389 AGATTAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1448
1776 AGATTTGTAATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1835
1449 ATAGGGTCTCTATCTTGGCATCTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1508
1836 TCCGTGTTCTCTATCTTGGCATCTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1895
1509 ATGAAACGCTTAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1537
1896 ACGAAACGTTGATTTCTGGAATGTTTC 1924

RESULT 11

US-10-128-714-38
; Sequence 38, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wendi
; APPLICANT: Tishkoff, Daniel

APPLICANT: Zamudio, Carlos
APPLICANT: Broshkin, Alexey M
APPLICANT: Lemieux, Sebastien M
TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
TITLE OF INVENTION: Methods of Use
FILE REFERENCE: 10182-018-999
CURRENT APPLICATION NUMBER: US/10/128,714
CURRENT FILING DATE: 2002-04-23
PRIOR APPLICATION NUMBER: US 60/285,697
PRIOR FILING DATE: 2001-04-23
PRIOR APPLICATION NUMBER: US 60/287,066
PRIOR FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/295,890
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: US 60/303,899
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: US 60/316,362
PRIOR FILING DATE: 2001-08-31
NUMBER OF SEQ ID NOS: 8603
SOFTWARE: PatentIn version 3.1
SEQ ID NO 38
LENGTH: 3398
TYPE: DNA
ORGANISM: Aspergillus fumigatus
US-10-128-714-38

Query Match 13.2%; Score 263.8; DB 15; Length 3398;

Best Local Similarity 56.3%; Pred No. 7.2e-56;
Matches 557; Conservative 0; Mismatches 372; Indels 60; Gaps 1;

QY 609 ATAGTCGGTTAAGGCTCTAGAGAGGTTCTCGATTCGCGCTTATAGGTTTGGATGCAAC 668
Db 1341 AGACTCCTCGCAAGAGCGCTCTAGCTTAATAAAGTAAAGGTTTCTCGACGCAAC 1400
QY 669 CTGCTTTGCAAGATGATTTTATCTGAATCTGTGATAGTGTCTTTCACAAATGTGTGG 728
Db 1401 CCGATTTGCAAGATGATTTTATCTGAATCTGTGATAGTGTCTTTCACAAATGTGTGG 1460
QY 729 CTGCTTTGTTGGTAACTGTCTATCTCTGGAATGCTTGTAGCAGCAAGGTTAACTAAAT 788
Db 1461 GTGTTGGCTTAGGCAATTCAGTGTATGTTGGAATTCGAATACCGGAGCGGTGACGAAAC 1520
QY 789 TATGTGATTTGGGGTGTGATGTTG----- 813
Db 1521 TTTGCAACTAAGAGATGATGATCTGTCAAGTGTGATGTTGATGTTGATGTTGATGAT 1580
QY 814 -----TGTTTGTCTGTTGTTGGGCTCAACGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 848
Db 1581 TTCTCTATATGTTGGCAATGATTTATGTTGTTTCTGATGATGATGATGATGATGATGAT 1640
QY 849 ATCTTGTCTGTTGGAATCAATGTTAAAGTTTCAAGTTTGGATGCGCAAGATGCAAGA 908
Db 1641 ATCTTTCAATAGGAATGCGCAAGGTTTGTACAGATATGGGATGAGAGCAATGTCGCC 1700
QY 909 AGATAAGATCAATGAGAGGCGCATCGTTACGTTGCGGGGCTTGGGCTGGAGTTTCACTCT 968
Db 1701 GTCTTGGCAATGATTTGGGCAATACCAATCGTGTGGGGCTCTTCTTGGAGCATGATA 1760
QY 969 TTTTGTCTTCTGGTGAAGGATTAAGATATTTATCAAGAGATATACGACAAAGAG 1028
Db 1761 TCCTTACGTCAGGTTCTCGGGATGCACTGATCTTCCACCGTGTGTTGTTTCCCGAGATC 1820
QY 1029 ATTTTGTGTAGTAACTGTCTGAGGACAAATCAGAGGTTTGTGACTGGAAGTGTCTATATG 1088
Db 1821 AGTATTTACGTCAGTGTCTCGGGCATAGCAGAGAGTTTGGGATCTCGGTGGACACCG 1880
QY 1089 ATAACCGTGAATGTTGAGGAAATGACAAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1148
Db 1881 AAGATGGCAACTAGCTTTCAGGCGGAAATGACAACTCATGTTGTTGTTGTTGTTGTTGTTGTT 1940
QY 1149 CAACCCGCTGTTCTCAAGTACTGTGAGCAACAAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1208
Db 1941 ATGAGACACCGCTTTATGCTTCTCGGACCATACCGGAGCGGTGAGGCGCATGCTGCTGTT 2000

Qy	1209	CTCCTCATCTTCATGGACCTCTTGTCATCTGGAGAGGAAGTCAGATAGATGTATTTCGTT	12668
Db	2001	CAACCCATCAACACCATCTCTCGCTCTAGGTGGAGCACGGCGGACCGGACCATCAAAAT	2060
Qy	1269	TTTGGAAATCAACCAAAATCTCACACTTAGCTGTATGGACACTGGAAAGTCAGGTTTGCA	1328
Db	2061	TCGTGGAACACTGCTACTCTGGTTCACTGATCAAGGAAGTTGATACCGGTAGCCAGTGTGTA	2120
Qy	1329	ATCTTGTCTGGTCCAAAATGTCAAAGAACTAGTAGCCAACATGGGTACTCCCGAAGCC	1388
Db	2121	ACCTGGCATGGTGGAGAACTCCGGACGAATCATCAGTACACACGGCTCAGTGCAGAAACC	2180
Qy	1389	AGATTATTGTTTGGAGATACCCCACTATGTCAAAGCTGGGACTCTTACCGGCACTACTT	1448
Db	2181	AGATTGAATCTGGAAATACCCGGCATGGAGCAGATTGTATCGCTACCGGCACACTT	2240
Qy	1449	ATAGGGTTCCTATCTTGCCATCTCTCCAGATGGACAGACTATTGTAACTGGAGCTGGAG	1508
Db	2241	TCGGTGTCTCTACCTTGGGATGAGCCCCGATGACAGACCGTCGTTACCGGAGCTGGTG	2300
Qy	1509	ATGAAACGGCTTAGGTCTGGAATGTTTTC	1537
Db	2301	ACGAAACGTTGAGATCTCTGGAAGATTTTC	2329

RESULT 12

```

US-10-128-714-5038
; Sequence 5038, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lenieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE OF INVENTION: Methods of Use
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5038
; LENGTH: 3993
; TYPE: DNA
; ORGANISM: Aspergillus fumigatus
US-10-128-714-5038

```

	Query Match	13.2%;	Score 263.8;	DB 15;	Length 3993;
	Best Local Similarity	56.3%;	Pred. No. 7.9e-56;		
	Matches 557;	Conservative 0;	Mismatches 372;	Indels 60;	Gaps 1;
Qy	609	ATAGTCCGGTTAAGCCTCCTAGGAAGTTCCTCGATCGCCTATAAGTGTGATGATCGAC	668		
Db	1936	AGACTCTCGCAAGCAGCCTCGCTACGTTAATAAAGTACCCTATAAGTTCTTCGACGCAC	1995		
Qy	669	CTGCTTTGCAAGATGATTTTTATCTGAATCTGGTAGATTGGTCTTCACACAATGTGTGG	728		
Db	1996	CCGATTTGCAAGATGATTTTTCTAGATCTGGTGGACTGGGCGAGTAGTAATGTGCTAG	2055		
Qy	729	CTGTTGGTTGGGTAACTGTGTCTATCTCTGGAATGCTTTAGCGACGAAGGTAACTAAAT	788		

RESULT 13

US-10-369-493-45642
 ; Sequence 45642, Application US/10369493
 ; Publication NO. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
 ; FILE REFERENCE: 38-10(52052) B

Db	2056	GTGTGGCCTAGCAATTCAGTGTACATGTGGAAATTCGAATACCGAGCGGTGACGAAC	2115
Qy	789	TATGTGATTTGGGGGTGTGATGTG	813
Db	2116	TTTGGCAACTAAGAGATGATATGTGTACCAAGTGTAGCTGGATTCAGAGGGTAATGAGAT	2175
Qy	814	-----TGTTTGTCTGTTGGTGGCTCAAGCTGTACTC	848
Db	2176	TTCTCTATATGTTGGCAATGAATATTATGTGTTCATGACTGACATTTGGACAGGGCAAC	2235
Qy	849	ATCTTGTCTGTGGAACTAAACAATGTGTAAAGTTTCAGATTGGGATGCGACAGATGCAAGA	908
Db	2236	ATCTTTCAATAGGAATCTGGCAAAAGTCTTTGTACAGATATGGATGCGACAGCATTTGCGCC	2295
Qy	909	AGATAAGATCAATGGAGGGCCATCGGTTCAGTGTTCGGGGCCCTTGGCCCTGGAGTTCACTC	968
Db	2296	GTCTTTCGGACAAATGATTTGGGCATACCAATCGTGTGGGGGCTCTTGTCTTTGGAAAGCATCAT	2355
Qy	969	TTTTTGTCTTCTGTGTGGACGGGATAAGAAATATTATTAACAAGAGATATACGCAACAAGAAG	1028
Db	2356	TCCTTACGTCAAGTTCTCGGATCGACTGATCTTCCACCGTGAATGTTGTTTGGAAATCAACACT	2415
Qy	1029	ATTTTGTGTAGTAACCTGTCAAGACACAAATCAGAGGTTTGTGGACTGAAGTGGTTCATATG	1088
Db	2416	AGTATTTTACGTTCGACTGTTCGGGCCATTAAGCGAGAAAGTTTTCGGGACTTCGGGTGGAAACACCG	2475
Qy	1089	ATAACCGTGAAGTTGGCATCTGGAGGAAATGACAAACAAATTTGTTTGTGGAAATCAACACT	1148
Db	2476	AAGATGGCCAACTAGCTTTCAGGCGGAAATGACAAACAACTCATGTGTGGACAAAGCTGA	2535
Qy	1149	CAACCCAGCCTGTCTCAGTACTGTGAGGACACAGCAGCTGTTTAAACTATTGCTATGGT	1208
Db	2536	ATGAGACACCGCTTTATCGGCTTCTCGGACCAATACCAGCAGCGGTGAAGGCCATTCGCTGT	2595
Qy	1209	CTCTCTATCTTCATGGACTTCTTGCACTCTGGAGGAGGAACCTGCAGATAGATGTATTCTTT	1268
Db	2596	CACCCCATCAACCACTCTCTCGCTCAGTGTGGAGCACGCGGACCGGACCACTCAAT	2655
Qy	1269	TTTGGATATCAACACAAACTCAACCTTAGCTGTATGGAACACTGGAAGTCAGGTTTGCA	1328
Db	2656	TCTGGAACACTGCTACTTGGTTCTCATCATCAAGGAAGTTGATACGGGTGACCGAGTGTGTA	2715
Qy	1329	ATCTTGTCTGTCCAAAATGTCAAACAACTAGTAGACACATGGGTACTCCACAGAAC	1388
Db	2716	ACCTGGCATGGTCGAAGAACTTCGACGMAATCATCAGTACACACGGCTACAGTCAGAAC	2775
Qy	1389	AGATTATTGTTGGAGATACCCCACTATATGTCAAAGCTGGCGACTCTTTACCGGCCATCTTT	1448
Db	2776	AGATTGTAAATCTGGAAATACCCGCGCATGAGCAGATTGTATCGCTGACCGGCCACATT	2835
Qy	1449	ATAGAGGTTCTCTATCTTTGCCATCTCTCCAGATGGACAGACTATTGTAACTGGAGCTGGAG	1508
Db	2836	TCCGTGTTCTCTACTCTTCGATGAGCCCGATGGACAGACCGTCTTACCGGAGCTGGTG	2895
Qy	1509	ATGAACCGCTTAGGTTCTTGGAAATGTTTTTC	1537
Db	2896	ACGAACCGTTTCAGATTCTTGGAAAGATTTTC	2924

QY	614	CCGGTTAAGCCTCTAGGAAGGTTCTCGATCGCCCTTATAGAGCTTTTGGATGCACCTGCT	673
DB	718	CCGGGGAAGCAATTTAGACAATAGCTAAAGTTCCCTATAGAGTTTTAGATGCCCATCA	777
QY	674	TTGCAAGATGATTTTATCTGAAATCTGGTAGATTGGTCTTTCACACAATGTGTGGCTGTT	733
DB	778	TTAGCAGACGATTTTACTACAGCTTGATAGATTGGTCAAGTACTGATGTTTGGCAGTT	837
QY	734	GGTTTGGGTAACTGTGCTATCTCTGGAATGCTTGTAGACGAAGGTAACTAAATATGTT	793
DB	838	GCTCTTGGAAAAATCGATTTTATTAACCGATAATAACACTGGCGACGTTGTTCACTATGC	897
QY	794	GATTTGGGGTGTGATGATTGTTGTTGTTCTGTGTTGGGCTCAACGTGTGTAATCATCTT	853
DB	898	GACACGGAAAAACGAATAACAAGCT-----TAAGCTGGATTGGAGCAGGCTCTCATCTG	951
QY	854	GCTGTTGGAACTAACAAATGTTAAAGTTTCAGATTTGGGATCGACCAAGATCAAGAAGATA	913
DB	952	GCAGTAGGTCAAGCAAAATGCACTGTAGAAATTTATACGTAATGAAAGAAATGATTT	1011
QY	914	AGATCAATAGAGGGCCATCGGTTACGTGTGCGGGCCTTGGCGCTGGAGTTCATCTCTTTTG	973
DB	1012	AGGACGTTGTGGGACATATTTGACAGATGACGTGTTTATCTTGGAAATAATCATGTTTGG	1071
QY	974	TCTTCTGCTGGACGGGATAGAAATATTTATCAACGAGATATACGACACAAGAGAAGATTTT	1033
DB	1072	ACGTCTGGGATAGAGATCATCGAATACTGCTATAGGGATGTTAGGATGCCAGATCCCTTT	1131
QY	1034	GTTAGTAAACGTGCAGGACACAAATCAGAGTTTGTGCACTGAAGTGGTCATATGATAAC	1093
DB	1132	TTTGAAACTATAGAATCGCACATCAGGAAGTCTGTGGCCCTAAAGTGGAAATGTACGGAC	1191
QY	1094	CGTGAGTTGGCATCTGGAGGAAATGACAAACAAATGTTTGTGTGGAAATCAACACTCAACC	1153
DB	1192	AACAAACTTGGCTCAGGTGGTAAACGATAACGCTGCCATGTTTACGAAGGAACGTCGAA	1251
QY	1154	CAGCGTGTCTCAAGTACTGTGAGCACACAGCAGCTGTTAAAGCTATTGCATGGTCTCCT	1213
DB	1252	TCTCCAATCTTGAAGTTGACGAACATAAGGCTGCTGTTAAGGCAATGGCCTGGTCTCCT	1311
QY	1214	CATCTTCAATGCACTTCTTGCATCTCGAGGAGGAACTGCAGATAGATGTATTGTTTTGG	1273
DB	1312	CATAAACGAGGTGTCTAGTACCAGTGGTGATACAGCAGATAGAAGCTAAAAATCTGG	1371
QY	1274	AATACACCAACAACTCACCTTAGCTGTATGAGCATCTGGAAGTCAAGTTTGCACCTT	1333
DB	1372	AACGTTAATACGTCAATAAAGATAGTGTATAGATACGGGCTCTCAATATGTAATATG	1431
QY	1334	GTCCTGGTCCAAAAATGTCAACGAACTAGTAAAGCACATGCGGTACTCCAGAACAGATT	1393
DB	1432	GTATGGTCAAGAAATACTAACGAGCTTGTGACATCACATGGTTACTCAAAAATAAATCTTA	1491
QY	1394	ATTGTTTGGAGATACCCCACTATGTCAAAGCTGGGCACTCTTACCGGCCATACTATAGG	1453
DB	1492	ACCTTGTGGGACTGTAACTCTATAGGATCCAAATTCGAATGCTCATAGTTTCAGG	1551
QY	1454	GTTCTCTATCTTGGCACTCTCTCCAGATGGACAGACTATGTAACTGGAGCTGGAGATGAA	1513


```
QY 1143 -----AACACTCAACCCAGCCTGTCTCAAGTACTGTGTGACACACAGCAGCTG 1190
Db 860 TCCTTCTCTCAAACTCAACACACAAATGGCTGCACAGGCTTGAGGACATACATCTGCTG 919
QY 1191 TTAAGCTATGTCATGGTCTCTCATCTTCATGGAATCTTGTGATCTGGAGGAGAACTG 1250
Db 920 TGAAGCTCTTGGTGGTGGCTTTCACAGCGAAATTTGCTTGAACCTGGTGGTGGAG 979
QY 1251 CAGATAGATGATTTGTTTTTGGAAATACACACAACTCAACCTTTAGCTGTATGGACA 1310
Db 980 GAGACAGGACGATAAGTTCTTGGAAATCTCAACTGGGGCTTGTGAATTCAGTAGACA 1039
QY 1311 CTGGAAGTCAAGTTTGAATTTGTCTGTGTCCTCAAAATGTCACGAACTAGTAAGCACAC 1370
Db 1040 CTGGTTCCCAAGTTTGTCTGTGTTATGGAGCAAGAAATGAAAGAGAGTTGCTTAGCTCAC 1099
QY 1371 ATGGGTACTCCAGAACCAATATTGTTTGGAGATACCCCACTATGTCAAAGCTGGCGA 1430
Db 1100 ATGGGTTTACAGAAATCAGCTTACACTTTGGAGTACCCATCCATCGGTGAAATGGCTG 1159
QY 1431 CTCTTACCGCCATCTATPAGGTTCTCTATCTTCCCATCTTCCAGATGGACAGACTA 1490
Db 1160 AGCTCACTGGTCTACATCAAGAGTTCTATATATGCGCCAGAGTCCAGATGGTTGACTG 1219
QY 1491 TTGTAAGTCAAGCTGGAGTGAACGCTTAGGTTCTGGATGTTTTT 1536
Db 1220 TAGCTTACGAGCAGGAGACGAGACCTGAGGTTTTTGGAACTGTTTT 1265
```

RESULT 15

```
US-09-938-842A-988
; Sequence 988 Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 988
; LENGTH: 1344
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-988
```

```
Query Match 10.6%; Score 213.6; DB 12; Length 1344;
Best Local Similarity 54.2%; Pred. No. 2.2e-43;
Matches 513; Conservative 0; Mismatches 409; Indels 24; Gaps 3;

QY 615 CGGTTAAGGCTCTAGGAAGGTTCTCGATCGGCTTATAAGGTTTGGATGCACTGCTT 674
Db 320 CTGTAAAGCCCTCGTCGATACATTCCTCAGACTTCTGAGAGAACCTTGGATGCACTGACA 379
QY 675 TGCAGATGATTTTATCTGAATCTGGTAGATGGTCTTACACAAATGTTGGCTGTTG 734
Db 380 TTGTTGACGATTTCTACTCTCACTTGTCTGACTGGGGAAGTGCATATGCTTAGCCATAG 439
QY 735 GTTTGGGTAATCTGTCTATCTCTGGAATGCTTGTAGCAGCAAGGTAATTAATATG 794
Db 440 CGTTGGACCACTGCTCTCTTGTGGGATGCTTCCACTGTTCTATCTGAGCTTGTGA 499
```

Search completed: January 23, 2004, 23:19:37

Job time : 721 secs

```
QY 795 ATTTGGGGTGTGATGATG---TGTGTTTCTGTGTTGGGCTCAACGTTGTTACTCATC 851
Db 500 CCAATTGATGAGGAGAAAGGACCTGTCAAAAGTATCAACTGGGCTCTGATGGTCTCATG 559
QY 852 TTGCTGTTGGNACTAAACAATGGTAAAGTTTCAAGTTTGGGA---TGCAAGCAAGATGCAAGA 908
Db 560 TTGCAGTTGGACTCAACAATCTGTAAGTCCAGCTGTGGGATTTCTGATCCAACCGTCAAC 619
QY 909 AGATAAGATCAATGGAGGGCCATCGTTAGCTGTGCGGGGCTTTGGCCCTGGAGTTCACTC 968
Db 620 TGAGAACATTAAAGGGTGGTCAACAGTCAAGAGTAGGATCACTGGCATGGAACAATCAT 679
QY 969 TTTTGTCTTCTGGTGGAGGGGATAAGATATTTATCAAGAGATATACGACATATACCAAGAAG 1028
Db 680 TCCTTACTACTGGAGGAATGGATGGACTGATCAATCAACAATGATGTGAGGATCAAGATCAC 739
QY 1029 ATTTTGTAGTAAACTGTCAAGGACACAAATCAGAGGTTTGTGGACTGAAGTGGTCATATG 1088
Db 740 CCAATTGTGGAACCTTACAGAGGTCACATCAAGAAGTTTGTGGCTCAAGTGGTCAGGAT 799
QY 1089 ATAAACGTTGATTTGGCATCTTGGAGGAAATGACAAATTTGTTTGTGTTGGAATC----- 1142
Db 800 CTGGACACAACTAGCAAGTGGTGGCAACGACAAATGTGGTACACATCTGGGATCGTTCTG 859
QY 1143 -----AACACTCAACCCAGCCTGTCTCTCAAGTACTGTGTGAGCACACAGCAGCTG 1190
Db 860 TCCTTCTCTCAAACTCAACACACAAATGGCTGCACAGCTTGAGGAACATACATCTGCTG 919
QY 1191 TTAAGCTATTTGATGGTCTCTCATCTTCAATGGAATCTTGTGATCTGGAGGAGAACTG 1250
Db 920 TGAAGGCTCTTGGCTGGTGGCTTTTCCAAAGCGAAATTTGCTTGCACACTGGTGGTGGAG 979
QY 1251 CAGATAGATGATTTGTTTTTGGAAATACAAACCAAACTCAACCTTAGCTGTATGGACA 1310
Db 980 GAGACAGGACGATAAGTTCTGGAATACTCACTGGGGCTTGTGAAATTCAGTAGACA 1039
QY 1311 CTGGAAGTCAAGTTTGCATCTTCTGTGGTCCAAAATGTCACGAACTAGTAAGCACAC 1370
Db 1040 CTGGTTCCCAAGTTTGTCTGTGTTATGGAGCAAGAAATGAAAGAGAGTTGCTTAGCTCAC 1099
QY 1371 ATGGGTACTCCAGAACCAAGATTTGTTTGGAGATACCCCACTATGTCAAAAGCTGGCGA 1430
Db 1100 ATGGGTTTACACAGAAATCAGCTTACACTTTGGAAGTACCCATCCATGATGTGAAATGGCTG 1159
QY 1431 CTCTTACCGGCCATCTATATAGGTTCTCTATCTTGGCATCTTCCAGATGGACAGACTA 1490
Db 1160 AGCTCACTGGTCTATACATCAAGAGTTCTATATATGCGCCAGAGTCCAGATGGTTGACTG 1219
QY 1491 TTGTAAGTCAAGCTGGAGATGAACGCTTAGGTTCTGGATGTTTTT 1536
Db 1220 TAGCTTACGAGCAGGAGACGAGACCTGAGGTTTTTGGAACTGTTTT 1265
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 23, 2004, 11:32:03 ; Search time 4356 Seconds
(without alignments)
11192.557 Million cell updates/sec

Title: US-09-701-572-1
Perfect score: 2006
Sequence: 1 gattcgccagcaggaagaa.....agaaaaaagaaaaa 2006

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estnu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_estc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_est3:*
- 12: gb_est4:*
- 13: gb_est5:*
- 14: gb_est6:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pln:*
- 20: em_gss_vrt:*
- 21: em_gss_fun:*
- 22: em_gss_mam:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_rod:*
- 26: em_gss_phg:*
- 27: em_gss_vrl:*
- 28: gb_gss1:*
- 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	667.2	33.3	759	10	BG584091
2	624.6	31.1	820	14	CA919262
3	620.6	30.9	630	12	BI270288
4	575.4	28.7	594	13	BQ146406

5	571.8	28.5	1633	11	AY112458
6	543.6	27.1	648	10	BG044933
7	499.4	24.9	677	13	BU042571
8	431.8	24.5	558	10	AW980816
9	468.2	23.3	987	10	BG444274
10	463.4	23.1	528	14	CB828401
11	456.8	22.8	680	14	CA837263
12	407	20.3	856	14	CB675428
13	392	19.5	710	9	AW030735
14	376.4	18.8	460	9	AW498616
15	342	17.0	909	14	CA764180
16	339.4	16.9	501	9	AI895812
17	331.2	16.5	1818	11	AK081476
18	327	16.3	577	14	CA020821
19	312.8	15.6	533	9	AW065517
20	311.8	15.5	2982	11	AK083686
21	303.6	15.1	350	9	AU252139
22	292.8	14.6	530	13	BQ754977
23	282.6	14.1	786	12	BI652526
24	281.2	14.0	557	10	BE329969
25	281.2	14.0	705	14	CD003448
26	281	14.0	913	12	BI904176
27	280.6	14.0	554	9	AI770538
28	278.8	13.9	570	14	CA638695
29	275	13.7	731	14	CB979301
30	273.6	13.6	873	12	BI153551
31	273.4	13.6	801	11	AY109717
32	273.2	13.6	591	10	BF003530
33	273.2	13.6	878	10	BE456106
34	270.6	13.5	571	12	BI420244
35	270.4	13.5	879	13	BQ962474
36	270.2	13.5	863	12	BI554814
37	269.2	13.4	614	14	CD003026
38	267.4	13.3	688	14	CA923074
39	265.6	13.2	469	12	BM500135
40	264.6	13.2	779	13	BQ572197
41	263.8	13.2	872	12	BI456001
42	263.2	13.1	647	14	CA925058
43	263.2	13.1	647	14	CA925755
44	262.2	13.1	848	12	BM945914
45	261.4	13.0	730	14	CB979227

ALIGNMENTS

RESULT 1
BG584091
LOCUS
DEFINITION
EST485851 MHAM Medicago truncatula/Glomsu versiforme mixed EST
Library cdna clone pMHAM-14B17 5' end, mRNA sequence.
ACCESSION
BG584091
VERSION
BG584091.1 GI:13599155
KEYWORDS
EST
SOURCE
Medicago truncatula/Glomsu versiforme mixed EST library
ORGANISM
Medicago truncatula/Glomsu versiforme mixed EST library
REFERENCE
1 (bases 1 to 759)
AUTHORS
Harrison,M.J., Liu,J., Town,C.D., Van Aken,S., Utterback,T., Cho,J.
and Frazer,C.M.
TITLE
ESTs from roots of Medicago truncatula after colonization with
Glomsu versiforme, 2001
JOURNAL
Unpublished
COMMENT
Contact: Harrison M.J.
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73401
Tel: 580-223-5810
Fax: 580-221-7380
Email: mharrison@noble.org
Noble EST name: N379561e TIGR sequence name: MTDBM09TK More
information is available at: http://www.medicago.org
Seq primer: SKmod (CTA GAA CTA gtg gat CC).

759 bp mRNA linear EST 11-APR-2001
EST485851 MHAM Medicago truncatula/Glomsu versiforme mixed EST
Library cdna clone pMHAM-14B17 5' end, mRNA sequence.


```

FEATURES
  source      Location/Qualifiers
              i. .759
              /organism="Medicago truncatula/Gloms versiforme mixed EST
              library"
              /mol_type="mRNA"
              /cultivar="Medicago truncatula genotype A17"
              /db_xref="taxon:119092"
              /clone="pVHAM-14817"
              /tissue_type="roots colonized with Gloms versiforme"
              /dev_stage="roots harvested at 10, 17, 22, 31 and 38 days
              post-inoculation with Gloms versiforme. The library was
              made from a mixture of RNA from each of these stages."
              /lab_host="E. coli strain XLOLR"
              /clone_lib="MHAM"
              /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
              XhoI; cDNA was prepared from polyA+ enriched RNA from
              roots harvested at 10, 17, 22, 31 and 38 days
              post-inoculation with Gloms versiforme. The cDNA was
              directionally ligated into the Unizap XR vector from
              Stratagene and packaged using Gigapack III Gold packaging
              extracts. Plasmids containing cDNA inserts were excised
              from the recombinant lambda-Zap phage using Ex-assist
              helper phage and propagated in XLOLR cells."
BASE COUNT   187 a 165 c 185 g 222 t
ORIGIN
Query Match      33.3%; Score 667.2; DB 10; Length 759;
Best Local Similarity 95.1%; Pred. No. 3.1e-78;
Matches 721; Conservative 0; Mismatches 33; Indels 4; Gaps 3;

Qy 109 ATATAAAGCGCTCAAAAATCTTTTACAGCGTCTTTTCCCGGGGAAAAAATTA 168
Db 1 ATATAAATCACTGAAGAGAAATCTTTTACAGCGTCTTTTCCCGGGGAAAAAATTAAC 60

Qy 169 ACACAGCTCCGCCATGACGGAACCGGTATCGAAATCCACACCGACTTCCACCGTCAG 228
Db 61 ACA-AGCTCCGCCATGACGGAACCGGTATCGAAATCCACACCGACTTCCACCGTCGG 119

Qy 229 AGACAAATCTCACCGCTGAGCACAACCGGAGAGTCTCGGTCATCTAGCGGTATGAT 288
Db 120 AGATAAATCTCCACCGCGGAGCCATCACCGGAGAGTCTCGGTCATCTAGCGGTATGAT 179

Qy 289 CAACAGCAACCATTAACACTCTCTCGAAACCAATCTACTCCGATAGGTTCAATCCGAG 348
Db 180 CAACAGCAACCATTAACACTCTCTCGAAACCAATCTACTCCGATAGGTTCAATCCGAG 239

Qy 349 TAGATCTGCTTCGAAATTCGCTTTGTTGATATCAATCTCCGACAGAGCGCGATGA 408
Db 240 TAGATCTGCTTCGAAATTCGCTTTGTTGATATCAATCTCCGACAGAGCGCGATGA 299

Qy 409 TAGTTCACGCGTTATACGACTCTCTCGAAACCGCGTTGTTGACCGGATGTTGCCGG 468
Db 300 TAGTTCACGCGTTATACGACTCTCTCGAAACCGCGTTGTTGACCGGATGTTGCCAGG 359

Qy 469 TCCGGTTACGCGGAAAAAACCAGCTCCCGTCGATGACATTCGCGAATAGGAATATTTT 528
Db 360 TCCGGTTACGCGGAAAAAACCAGCTCCCGTCGATGACATTCGCGAATAGGAATATTTT 419

Qy 529 TAGGTATAGAGGAGGAGAGACAGTCCATCGCTTTCGCGGTTTATGGATGATGA 588
Db 420 TAGGTATAGAGGAGGAGAGACAGTCCATCGCTTTCGCGGTTTATGGATGATGA 479

Qy 589 TTTTGTTCCTGGTGTAAATCATAGTCCCGTTAAGGCTCCTAGGAAGGTTCTTCGATGCC 648
Db 480 TTTTGTTCCTGGTGTAAATCATAGTCCCGTTAAGGCTCCTAGGAAGGTTCTTCGATGCC 539

Qy 649 TTATAAGTTTGGATGACCTGCTTCGACATGATTTTATCTGAATCTGGTAGATTG 708
Db 540 TTATAAGTTTGGATGACCTGCTTCGACATGATTTTATCTGAATCTGGTAGATTG 599

Qy 709 GTCTTCACACAAATGTGTGGCTGTGGTTTGGGTAACGTGTCTATCTCTGGAATGCTTG 768
Db 600 GTCTTCACACAAATGTGTGGCTGTGGTTTGGGTAACGTGTCTATCTCTGGAATGCTTG 659

Qy 769 TAGCAGCAGAGTAACTAAATTAATCTGATT-TGGGGGTTGATGATTGTG--TTTGTCTGT 825
Db 660 TAGCAGCAGAGTAACTAAATTAATTAATCTGATTGTTGGGGGTTGATGATTGTTGTCTGT 719

Qy 826 TGGTTGGGCTCAACGCTGCTACTCATCTTCTGCTTGGAA 863
Db 720 TGGTTGGGCTCAACGCGGTACTCATCTTCTGCTTGGAA 757

RESULT 2
CA919262/c 820 bp mRNA linear EST 27-DEC-2002
LOCUS EST635980 MTUS Medicago truncatula cDNA clone MTUS-11B5, mRNA
DEFINITION sequence.
ACCESSION CA919262
VERSION CA919262.1 GI:27406192
KEYWORDS EST.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosoids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
1 (bases 1 to 820)
VandenBosch,K., Endre,G., Silverstein,K. Town,C.D., Van Aken,S.,
Utterback,T., Cheung,F. and Fraser,C.M.
The Medicago truncatula 6K unigene set: cDNA clones selected and
re-arrayed from various libraries
Unpublished
Contact: VandenBosch K
Department of Plant Biology
University of Minnesota
220 Biosci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
Tel: 612 624 2755
Fax: 612 625 1738
Email: kvandenb@cs.umn.edu
TIGR sequence name: MTUAK17V
More information is available at: www.medicago.org
Seq primer: gta ata cga ctc act ata ggg c).
FEATURES
  source      Location/Qualifiers
              1..820
              /organism="Medicago truncatula"
              /mol_type="mRNA"
              /cultivar="A17"
              /db_xref="taxon:3880"
              /clone="MTUS-11B5"
              /tissue_type="mixed tissues"
              /dev_stage="various stages"
              /lab_host="XLOLR"
              /clone_lib="MTUS"
              /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
              XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
              was directionally ligated into the Unizap XR vector from
              Stratagene and packaged using Gigapack III Gold packaging
              extracts. Plasmids containing cDNA inserts were excised
              from the recombinant lambda-Zap phage using Ex-assist
              helper phage and propagated in XLOLR cells."
BASE COUNT   271 a 155 c 161 g 233 t
ORIGIN
Query Match      31.1%; Score 624.6; DB 14; Length 820;
Best Local Similarity 94.9%; Pred. No. 1.1e-72;
Matches 701; Conservative 0; Mismatches 29; Indels 9; Gaps 5;

Qy 1272 GGAATACACCAAACTCACACCTTAGCTTAGTATGACACTGAGTTCAGGTTGCAATC 1331
Db 820 GGAATACACCAAACTCACACCTTAGCTTAGTATGACACTGAGTTCAGGTTGCAATC 762

Qy 1332 TTTCTCTGGTCCAAAATGTCAACGAATAGTAGACACATGGGTACTCCAGAACCA 1391
Db 761 TTTCTCTGGTCCAAAATGTCAACGAATAGTAGACACATGGGTACTCCAGAACCA 702

```


1392 TTATTGTTGGATACCCCACTATGTCAAAGCTGGCACTCTTACCGCCATCTTATA 1451
Db
701 TAATTGTTGGATACCCCACTATGTCAAAGCTGGCACTCTTACCGCCATCTTATA 642
Qy 1452 GGGTTCTCTATCTTGGCACTCTTCCAGATGGAGCACTATTTAATCTGGAGCTGGAGATG 1511
Db 641 GGGTTCTCTATCTTGGCACTCTTCCAGATGGAGCACTATTTAATCTGGAGCTGGAGATG 582
Qy 1512 AAACGCTTAGTCTTGGCACTCTTCCCTTCCCTTAAATCAACAGATCTGAAAGTGAAA 1571
Db 581 AAACGCTTAGTCTTGGCACTCTTCCCTTCCCTTAAATCAACAGATCTGAAAGTGAAA 522
Qy 1572 TCGAGCATATCTCTTGAAGAACTACTATACAGTGATGATCTTCCGCTGGAGCCCA 1631
Db 521 TCGAGCATATCTCTTGAAGAACTACTATACAGTGATGATCTTCCGCTGGAGCCCA 462
Qy 1632 ATCATGTGGCATATTTCTAAGTTTGGTGTCTGTGAGAACTAAATTTCTGAGCGAGAA 1691
Db 461 ATCATGTGGCATATTTCTAAGTTTGGTGTCTGTGAGAACTAAATTTCTGAGCGAGAA 402
Qy 1692 CACCATGGTGGAAAAACCTTGAATATAAAACACACCAAGTAGCATCTTTACCAACTG 1751
Db 401 CACCATGGTGGAAAAACCTTGAATATAAAACACACCAAGTAGCATCTTTACCAACTG 342
Qy 1752 GGAGAGCTTGGAGGAGCTATAAAGTTTGTATATGGCTGCGGTGA---TATTCCCTG 1807
Db 341 GGAGAGCTTGGAGGAGCTATAAAGTTTGTATATGGCTGCGGTGA---TATTCCCTG 282
Qy 1808 CATTATGTAGTCTCTATTTATTTAAGAAAGATGATACCAATGGTAAATTTATGTC 1867
Db 281 CATTATGTAGTCTCTATTTATTTAAGAAAGATGATACCAATGGTAAATTTATGTC 222
Qy 1868 TTGAGCTTATAC-ATGCAATGATGGAGTTGATGACCAAGTTTATTTATCTCTTTTTC 1926
Db 221 TTGAGCTTATACATGATGATGGAGTTGATGACCAAGTTTATTTATCTCTTTTTC- 163
Qy 1927 TTTCTCTTTTGTATAGTCTCTCTGCAATTTATATATTTAAGATGCGTTAAACAG 1986
Db 162 TTTCTCTTTTGTATAGTCTCTCTGCAATTTATATATTTAAGATGCGTTAAACAG 105
Qy 1987 AGAAAAAATAAAAAA 2005
Db 104 AGAAAAATCTCTTAACAAA 86

RESULT 3
BI270288
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BI270288 630 bp mRNA linear EST 18-JUL-2001
NF006A08FL1F1065 Developing flower Medicago truncatula cDNA clone
BI270288
BI270288.1 GI:14877742
EST
Medicago truncatula (barrel medic)
Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; rosids
; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.

1 (bases 1 to 630)
Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.
Expressed Sequence Tags from the Samuel Roberts Noble Foundation
Medicago truncatula flower library
Unpublished
Contact: May GD

Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7391
Fax: 580 221 7380
Email: gdmay@noble.org
Insert Length: 630 Std Error: 0.00

Plate: 006 row: A column: 08
Seq primer: TCACAGAGAAACAGCTATGAC.
Location/Qualifiers
1. -630
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF006A08FL1"
/tissue_type="Developing flowers"
/dev_stage="Developmentally pooled. Contains a mixture of
very young, developing, fully-opened flowers and flowers
in early transition into pods."
/clone_lib="Developing flower"
/notes="Vector: lambda Zap, cDNA was prepared from polyA+
enriched, pooled samples of equivalent amounts of total
RNA from very young, developing, fully-opened flowers and
flowers transitioning into pods. The cDNA was
directionally ligated into the Uni-Zap XR vector
(Stratagene) and packaged using the Gigapack III Gold
packaging extracts. Phagemids containing Uni-Zap XR vector
in vivo excised from the recombinant Uni-Zap XR vector
using EX881st helper phage and the E. coli strain
XL1-Blue MRF' (Stratagene). Excised plasmids were plated
using SOLR cells."

BASE COUNT 159 a 167 g 204 t 3 others
ORIGIN
Query Match 30.9%; Score 620.6; DB 12; Length 630;
Best Local Similarity 98.9%; Pred. No. 4.4e-72;
Matches 623; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 513 CGAATAGGAATATTTTAGGTATAGAGCGAGAGACAGATCCATCGCTTTCGC 572
Db 1 CGAATAGGAATATTTTAGGTATAGAGCGAGAGACAGATCCATCGCTTTCGC 60
Qy 573 CGTTATGGATGATGTTTCTCGTGTATATCATAGTCCGGTTAAGCTCCTAGGA 632
Db 61 CGTTATGGATGATGTTTCTCGTGTATATCATAGTCCGGTTAAGCTCCTAGGA 120
Qy 633 AGGTTCTCGATCGCTTATAAGGTTTGGATGACCTGCTTTCAGAGATGATTTTATC 692
Db 121 AGGTTCTCGATCGCTTATAAGGTTTGGATGACCTGCTTTCAGAGATGATTTTATC 180
Qy 693 TGAATCTGGTAGATGGTCTTCAACAATGTTGGCTGTTGGTAACTGTTGTT 240
Db 181 TGAATCTGGTAGATGGTCTTCAACAATGTTGGCTGTTGGTAACTGTTGTT 240
Qy 753 ATCTCTGAATGCTTGTAGCAGCAAGGTAACTAAATATGATGATTTGGGGTTCATGATT 812
Db 241 ATCTCTGAATGCTTGTAGCAGCAAGGTAACTAAATATGATGATTTGGGGTTCATGATT 300
Qy 813 GTGTTTGTCTGTGGTGGCTCAACGCTGATCTCATCTTGTCTGTGGTAACTAACATG 872
Db 301 GTGTTTGTCTGTGGTGGCTCAACGCTGATCTCATCTTGTCTGTGGTAACTAACATG 360
Qy 873 GTAAAGTTTCAGATTTGGGATGACGACAGATGCAAGAGATCAATGAGGCGCATC 932
Db 361 GTAAAGTTTCAGATTTGGGATGACGACAGATGCAAGAGATCAATGAGGCGCATC 420
Qy 933 GTTACGTTCTGGGCGCTTGGCTGGAGTTTCATCTCTTTTCTTCTGTGGTAACTAACATG 992
Db 421 GTTACGTTCTGGTGGCTTGGCTGGAGTTTCATCTCTTTTCTTCTGTGGTAACTAACATG 480
Qy 993 AGAATATTTATCAACGAGATATAGCAGCAAGAGATTTTGTAGTAACTGTCAGAC 1052
Db 481 AGAATATTTATCAACGAGATATAGCAGCAAGAGATTTTGTAGTAACTGTCAGAC 540
Qy 1053 ACAATATCAGAGTTTGTGGACTGAAAGTGGTTCATATGATAACCGGAGTTGCACTCGAG 1112
Db 541 ACAATATCAGAGTTTGTGGACTGAAAGTGGTTCATATGATAACCGGAGTTGCACTCGAG 600
Qy 1113 GAAATGACACAAATTTGTTTGGATTC 1142


```

Db      601 GAATGACACAAATTGTTTGGTATC 630

RESULT 4
B0146406
LOCUS   NF069G04FL1036 Developing flower Medicago truncatula cDNA clone
DEFINITION
ACCESSION B0146406
VERSION   NF069G04FL 5', mRNA sequence.
KEYWORDS EST.
ORGANISM Medicago truncatula (barrel medic)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 594)
AUTHORS Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.
TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation
JOURNAL Medicago truncatula flower library
COMMENT Unpublished
Contact: May GD
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7391
Fax: 580 221 7380
Email: gdmay@noble.org
Insert Length: 594 Std Error: 0.00
Plate: 069 row: G column: 04
Seq primer: TCACACGAGAACAGCTATGAC.
Location/Qualifiers
1..594
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF069G04FL"
/tissue_type="Developing flowers"
/dev_stages="Developmentally pooled. Contains a mixture of
very young, developing, fully-opened flowers and flowers
in early transition into pods."
/clone_lib="Developing flower"
/note="Vector: Lambda Zap; cDNA was prepared from polyA+
RNA from very young, developing, fully-opened flowers and
flowers transitioning into pods. The cDNA was
directionally ligated into the Uni-Zap XR vector
(Stratagene) and packaged using the Gigapack III Gold
packaging extracts. Phagemids containing cDNA inserts were
in vivo excised from the recombinant Uni-Zap XR vector
using ExAssist helper phage and the E. coli strain
Xli-Blue MRP (Stratagene). Excised plasmids were plated
using SOLR cells."
BASE COUNT 141 a 144 c 141 g 167 t 1 others
ORIGIN

Query Match 28.7%; Score 575.4; DB 13; Length 594;
Best Local Similarity 98.0%; Fred. No. 3.8e-66;
Matches 582; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 170 CACAGCTCCGATCGGACGACCGGTAAATCGAATCCACACCGACTCCACCGTCGAGA 229
DB 1 CACAGCTCCGATCGGACGACCGGTAAATCGAATCCACACCGACTCCACCGTCGAGA 60
QY 230 GACATTCTCCACCGCTGAGCCATCAGCGGAGTCTCCGTATGAGCGGTATGATC 289
DB 61 GATAATTCTCCACCGCGGAGCCATCAGCGGAGTCTCCGCAATGTAAGCGGTATGATC 120
QY 290 AACAGCAACATTACACCTTCACTTCTCGAACAAATCTACTCCGATAGTTCATTGGAGT 349
DB 121 AACAGCAACATTACACCTTCACTTCTCGAACAAATCTACTCCGATAGTTCATTCCGAGT 180

```

```

QY 350 AGATCTGCTTCGAATTCGCTTTGTTGATATCAATCTCCGACAGAGGACGGATGAT 409
DB 181 AGATCTGCTTCGAATTCGCTTTGTTGATATCAATCTCCGACAGAGGACGGATGAT 240
QY 410 AGTTCCAGCGCTTATACGACTCTTCTGAGAACGGCGTTGTTGGACCGGATGTTGCCGGT 469
DB 241 AGTTCACGCGCTTATACGACTCTTCTGAGAACGGCGTTGTTGGACCGGATGTTGCCAGT 300
QY 470 CGGTTACGCGGAAAACCGACTCCGCGTCGATGACATTCGCGAATAGGAATATTTTT 529
DB 301 CCGGTTACGCGGAAAACCGACTCCGCGTCGATGACATTCGCGAATAGGAATATTTTT 360
QY 530 AGGTATAACACGACGACGACGACGACGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 589
DB 361 AGGTATAACACGACGACGACGACGACGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 590 TTGTTCTGCTGGTAAATCATAGTCCGTTAAGGCTCTAGGAGGTTCTTCGATCGCT 649
DB 421 TTGTTCTGCTGGTAAATCATAGTCCGTTAAGGCTCTAGGAGGTTCTTCGATCGCT 480
QY 650 TATAAGGTTTTCGATGCACCTGCTTTGCAAGATGATTTTATCTCAATCTGCTAGATTGG 709
DB 481 TATAAGGTTTTCGATGCACCTGCTTTGCAAGATGATTTTATCTCAATCTGCTAGATTGG 540
QY 710 TCTTCACACATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 763
DB 541 TCTTCACACATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 594

```

```

RESULT 5
LOCUS AY112458 1633 bp' mRNA linear HTC 17-OCT-2002
DEFINITION Zea mays CL250.1 mRNA sequence.
ACCESSION AY112458
VERSION AY112458.1 GI:21217048
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 1633)
AUTHORS Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,
Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 1633)
AUTHORS Coe, E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.

```

```

FEATURES
source
1..1633
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:631015"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed Dupont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize

```



```

BASE COUNT      389 a      356 c      386 g      391 t      111 others
ORIGIN

Query Match      28.5%; Score 571.8; DB 11; Length 1633;
Best Local Similarity 72.8%; Pred. No. 5.7e-66;
Matches 724; Conservative 0; Mismatches 264; Indels 6; Gaps 1;

QY 622 GGCTCCCTAGGAAGGTTCTCGATCGCTTATAAGGTTTTCGATGCACTGCTGTTGCAAGA 681
DB 31 GGGMCCAGAGAGTCCCGAGGTCGCCCTACAGGTGCTGATGCGCTGCGTTGGAGA 90
QY 682 TGATTTTATCTGAATCTGTAGATGGTCTTACACAATGTGTGCTGTTGGTGG 741
DB 91 CGACTTCTACTCAACCTCGTCGATGGTCTTACATACGCTCGCTGTTGGTGG 150
QY 742 TAACGTGTCTATCTCGAATGCTTGTACGACGAGTAACTAAATATGTGATTTGG 801
DB 151 CAATGCGTTTACCTGTGGANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 210
QY 802 GGTTCATGATTTGTTGTTGTTGTTGCTCAACGTTGCTCACTCATCTTCTGTTGG 861
DB 211 NNNNNNNNCAACGTTTGTGTTGCTGGCGCAGCGGCACTCACCTTGCAGTAGG 270
QY 862 AACTAACTAGTAAAGTTTCAAGTTTGGGATGAGCAGATGCAAGATATAGATCAAT 921
DB 271 AACAAATCAAGAAAGTTTCAAGTCTGGGATGCAACTCGCTGTAAAGAAATAAGAACTAT 330
QY 922 GGAGGGCATCGTTAGCTGTGCGGGGCTTGGCGTCAACGTTGCTATCTCTTTGTTCTG 981
DB 331 GGAAAGCCATCGCATGCTGTAGGTGCTTTGATGAGTCTTCTGTTGTTCTTCTG 390
QY 982 TGGACGGATAGATATTTATCAACAGATATAGCACACAGAGATTTTCTTAGTAA 1041
DB 391 AAGCGTGCACAGAGATCTCCACCATGATATCCGCTCAAGAGATCATGTTAGCAA 450
QY 1042 ACTGTCAAGACACAAATCAGAGGTTTGTGACTGAAGTGTCTATATATCAACCGTGAGTT 1101
DB 451 GCTTACTGGGATATTAATCTGAGTTTGTGACTCAAGTGTCTTATGACAACTGTCAGCT 510
QY 1102 GGATCTGGAGAAATGACAAATTTGTTGTTTGAATCAACACTCAACCCAGCTGT 1161
DB 511 TGCATCGGTTGCAATGACACAGACTTTTGTGTTTGAATCCATTCAGTACAGCCAGT 570
QY 1162 CCTCAAGTACTGTGAGCACACAGCAGCTGTTAAAGTATTGATGCTCTCTCATCTTCA 1221
DB 571 ACTGAGTATACTGAGCACACAGCAGCTGTCAAGACTATTGCTGTGTCACCTCATCTCCA 630
QY 1222 TGGACTTCTTCATCTGAGGAGAACTGACATAGATAGATGTTGTTTGGAAATACAC 1281
DB 631 TGGGCTGCTTCATCTGTGAGGAACTGACATAGATGATGATGCAATTTTGGAAATACGAC 690
QY 1282 CACAACTCACACTTACTGTATGACACTGGAAGTCAAGTTTGCATCTTGTCTGCTC 1341
DB 691 CACAAATGCACACTTACTGTGAGCAGCTGAGCAGCAGCTTGCATCTTGTGCTGCTC 750
QY 1342 CAAAAATGTCACAACTAGTAAAGCACACATGSGTACTCCAGAACACAGATTTGTTTG 1401
DB 751 AAAGAATGTAATGAGCTTGTAGCACCCAGGCTACTCTCAGAACCAATAATTTGTTG 810
QY 1402 GAGATACCCCATGATCAAGCTGGGACCTTTACCGGCATCTATATAGGTTCTCTA 1461
DB 811 GAGATATCCAAATGTCGAGCTTGCACCTTTGACAGGGCATATCATACAGATTTGTA 870
QY 1462 TCTTCCCATCTCTCAGATGACACACTATTTGTAACCTGGAGCTGGAGATGAAGCGTTAG 1521
DB 871 CTTAGCTATTTCCCTGATGCCAGACCATAGTTACTGGTCTGTGTGATGAACGCTCG 930
QY 1522 GTTCTGGAATGTTTCCCTTCCCTTAATCAAGATATCAAGATCTGAAGTGAAG 1575
DB 931 GTTTTGGAAAGCTGTTTCCCTCTCCAAAGTCCAGAGTTCTGACAGTTTAAAGTTCGGT 990
QY 1576 AGCATTTATCTTCTGGAAGAACTACTATCATCAGTGA 1609

RESULT 6
BG044933
LOCUS
DEFINITION
  ssa35d02.y1 Gm-cl059 Glycine max cDNA clone GENOME SYSTEMS CLONE
  ID: Gm-cl059-939 5' similar to TR:Q9XED5 Q9XED5 CELL CYCLE SWITCH
  PROTEIN. ; mRNA sequence.
ACCESSION
  BG044933
VERSION
  BG044933.1 GI:12492159
KEYWORDS
  EST.
SOURCE
  Glycine max (soybean)
ORGANISM
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
  ; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
  Glycine.
REFERENCE
  1 (bases 1 to 648)
  Shoemaker R., Kelm, P., Vodkin, L., Erpelding, J., Coryell, V., Khanna
  A., Bolla, B., Warra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
  Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers
  Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk
  R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
  R., Waterston, R. and Wilson, R.
  Public Soybean EST Project
  Unpublished
  Contact: Shoemaker R/Public Soybean EST Project
  Public Soybean EST Project
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: eschwats@wustl.edu
  This clone is available through: ResGen, Invitrogen Corp. 2130
  South Memorial Parkway Huntsville, AL 35801 For further information
  call: (800)-533-4363 or contact via email: ccu@resgen.com
  High quality sequence stop: 405.
  Location/Qualifiers
    1..648
    /organism="Glycine max"
    /mol_type="mRNA"
    /db_xref="taxon:3847"
    /clone="GENOME SYSTEMS CLONE ID: Gm-cl059-939"
    /tissue_type="whole seedling, 2 week old, etiolated"
    /lab_host="DH10B"
    /clone_lib="Gm-cl059"
    /note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
    XhoI; The cDNA library was constructed from mRNA isolated
    from 2 week old etiolated whole seedlings of PI468916.
    Complementary DNA was synthesized from mRNA using a primer
    consisting of a poly(dT) sequence with a XhoI restriction
    site. EcoRI adapters were ligated to the blunt-ended cDNA
    fragments followed by XhoI digestion. The cDNA fragments
    were directionally cloned into the EcoRI- XhoI restriction
    site of the pBluescript vector. The ligated cDNA fragments
    were transformed into DH10B host cells (Gibco BRL). This
    library was constructed in the laboratory of Dr. Randy
    Shoemaker at Iowa state university."
  BASE COUNT      173 a      132 c      161 g      181 t
  ORIGIN

Query Match      27.1%; Score 543.6; DB 10; Length 648;
Best Local Similarity 90.0%; Pred. No. 5.2e-62;
Matches 582; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 860 GGAACCTAACATGTTAAAGTTTCAGATTTCGGGATGCAAGATGCAAGAGATAGATCA 919
DB 1 GGAACCTAGCATGTTAAAGTTTCAGATTTCGGGATGCAATCTCGATGCAAGAGATAGATCT 60
QY 920 ATGGAGGGCCATCGGTTACGTTGCGGGCCTTGGCGCTGGAGTTCACTCTTTTGTCTTCT 979
```


Dd		61	AUGGAGGGTCATCGGTTA	CGGGTTACGGGCTTGGGCTCTGGCTTGGAGTGTCATCTCTTTTGTCTCTCT	120
Qy		980	GGTGGACGGGATAAGAATA	TATTATCAACGAGATATACGCACAACAAGAAGATTTTGTTAGT	1039
Dd		121	GGCGGCAGGGATAAGAATA	TTTTATCAAAGAGATATCCGTGCACAAGAAGATTTGTCACT	180
Qy		1040	AAACTGTCAGGACACA	AAATCAGAGGTTTGTGGACTGAAGTGGTCATATGATGATAACCGTGAG	1099
Dd		181	AAATTGTCAAGGCCA	AAATCAGAGGTTTGTGGACTGAAGTGGTCTTATGATAACCGTGAG	240
Qy		1100	TTGGCATCTCTGGAGGA	AAATGACAATTTGTTTGTTTTGGAAATCAACACTCAACCCCAGCCT	1159
Dd		241	TTGGCATCTCTGGAGGA	AAATGACAATTTGTTTGTTTTGGAAATCAACACTCAACCTCACGCT	300
Qy		1160	GTCCCTCAAGTACTGTG	AGCACACAGCAGCTGTTAAAGCTATTGTCATGGTCTCTCTCATCTT	1219
Dd		301	GTCCCTGAAGTACTGTG	AGCATACAGCAGCTGTTAAAGCTATTGCAATGGTCTCTCTCATCTT	360
Qy		1220	CATGGCACTTCTGC	ATCTGGAGGAGGAATGCGAGATAGATGTATTCGTTTTTGGAAATACA	1279
Dd		361	CACGGACTTCTTG	CATCTGGGAGGAACATGCGACCGATGCATACGTTTTCTGGAAATACA	420
Qy		1280	ACCACAACTCAC	ACCTTACTGTTATGGACACTGGAAGTCAGGTTTGGCAATCTTGCTCG	1339
Dd		421	ACCACAACTCAC	ATTTAAGCTGATGGACACGGGAAGTCAGGTTTGGCAATCTTGCTCG	480
Qy		1340	TCCAAAANAATGTC	AAACGAACTAGTAAGCACCATGCGGTACTCCACGAACCAAGATTAATGTT	1399
Dd		481	TCCAAAANAATGTC	AAATGAAC TAGTAAGCACCATGGCTATTCCACGAACCAAGATTAATGTT	540
Qy		1400	TGGAGATACCCCA	CTATGTCAAAGCTCGGCAGCTCTTACCGGCCATCTTATAGGTTCTC	1459
Dd		541	TGGAGATACCCCT	CCATGTCAAAGTTGGCCACTCTTACGGGTCACTACTACCGAGTCTT	600
Qy		1460	TATCTTGCCCATCT	CTCCAGATGGACAGACTATTGTAATCGAGCTGG	1506
Dd		601	TATCTTGCCCATG	CTCCGATGGCAGACTATTGTAACCTGAGCTGG	647

RESULT	7	BU042571	677 bp	mRNA	linear	EST 26-AUG-2002
LOCUS						
DEFINITION		PP LEA0013C06f Peach developing fruit mesocarp Prunus persica cDNA clone PP LEA0013C06f, mRNA sequence.				

FEATURES	source
high quality sequence spot: 677.	
Location/Qualifiers	
1. .677	
/organism="Prunus persica"	
/mol type="mRNA"	
/cultivar="Loring"	
/db xref="taxon.3760"	
/cdone="bp.13a013206f"	
RESULT 8	
AW980816	
LOCUS	
EST391989	558 bp mRNA linear EST 07-SEP-2000
DEFINITION	GVN Medicago truncatula cDNA clone pGVN-58F10, mRNA
ACCESSION	AW980816
	sequence.

700CTV0898FF FF=SMTCZ/

```

VERSION
KEYWORDS
SOURCE
ORGANISM

AW980816.1 GI:8172357
Medicago truncatula (barrel medic)
Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids 1; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE
1 (bases 1 to 558)
AUTHORS
Fedorova,M., Pierson,B.L., Samac,D.A., Vance,C.P., Gantt,J.S., Peng
,H., Ellis,L., Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S.,
Holt,I.E. and Fraser,C.M.
TITLE
ESTs from one month old nitrogen-fixing root nodules of Medicago
truncatula
JOURNAL
Unpublished
COMMENT
Contact: Carroll P. Vance
Department of Agronomy and Plant Genetics
University of Minnesota
411 Borlaug Hall, 1991 Upper Buford Circle, St.Paul, MN 55108 USA
Tel: 612 625 5715
Fax: 651-649-5058
Email: vance004@maroon.tc.umn.edu
University of Minnesota data:M256650e
TIGR sequence name:MTCEB29TK
More information is available at: . http://chryslie.tamu.edu/medicago
Seq primer: SKmod (CTA GAA CTA gta gAT CC).
FEATURES
Location/Qualifiers
source
1..558
/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="pGVN-58P10"
/tissue_type="N2-fixing root nodules"
/dev_stage="effective root nodules harvested one month
post inoculation with Sinorhizobium meliloti"
/lab_host="E. coli strain XL0LR"
/clone_lib="GVN"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; cDNA was prepared from polyA+ enriched RNA from
effective root nodules harvested one month post
inoculation with Sinorhizobium meliloti. The cDNA was
directionally ligated into the Uni-ZAP XR vector from
Stratagene and packaged using Gigapack III Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant lambda-ZAP phage using Ex-Assist
helper phage and propagated in XL0LR cells."
BASE COUNT
172 a 112 c 120 g 154 t
ORIGIN
Query Match 24.5%; Score 491.8; DB 10; Length 558;
Best Local Similarity 93.3%; Pred. No. 3,4e-55;
Matches 514; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
QY 1260 GTATTGCTTTTGGAAATACACCAACTCACACCTTAGCTGTATGGACACTGGAAGTC 1319
Db 1 GTATTGCTTTTGGAAATACACCAACTCACACCTTAGCTGTATGGACACTGGAAGTC 60
QY 1320 AGTTTGCATCTTGTCTGTTCCAAAATGTCAACGAAGTGTAGTAAGCACACATGGGTACT 1379
Db 61 AGTTTGCATCTTGTCTGTTCCAAAATGTGGATGAAGTGTAGTAAGCACACATGGGTACT 120
QY 1380 CCCAGACCCAGATTTGTTGGAGATACCCCATGTATGCAAGCTGGCGACTTACCG 1439
Db 121 CCCAGACCCAGATTTGTTGGAGATACCCCATGTATGCAAGCTGGCGACTTACCG 180
QY 1440 GCCATACCTTATAGGTTTCTTATCTTGCCATCTCTCCAGATGGACAGACTATTGTAACTG 1499
Db 181 GCCATACCTTATAGGTTTCTTATCTTGCCATCTCTCCAGATGGACAGACTATTGTAACTG 240
QY 1500 GAGCTGGAGATGAACCGTATAGTTCTGGAAATGTTTCCCTCCCTTAATCAAGATA 1559
Db 241 GAGCTGGAGATGAACCGTATAGTTCTGGAAATGTTTCCCTCCCTTAATCAAGATA 300

1560 CTGAAAGTGAATCGAGCATTTATCTCTTGGAGAACTACTATCAGGTGATGATCTCTGG 1619
301 CTGAAAGTGAATCGAGCATTTATCTCTTGGAGAACTACTATCAGGTGATGATCTCTGG 360
1620 CGTTGCAGCCCAATCATGTGTCATATTTCTTAAGTTTGGTTGCTGTGTAGAACTAAATTT 1679
361 CATTCGACCCCTATCATGTGTCATATTTCTTAAGATTGGAGAGATAGTAGTAACTAATTA 420
1680 CTGACGGGAGAACACCATGTGGGAAAAACCTTGAATATAAAACACCAACCAAGTAGCAT 1739
421 TATAGCATAGAACACCATGTGGAGATAAACCTAGAGTATATAAAACACCAACCAAGTAGCAT 480
1740 CTTTACCACCTGGGAGAGCCCTTGGGGAGAGCTATAAAGTTTTCATATGCTGCCGGTGA 1799
481 CTTTACCACCTGGGAGAGCCCTTGGGGAGAGCTATAAAGTTTTCATATGCTGCCGGTGA 540
1800 TATTCTCTGCAT 1810
541 TATACTTTTAT 551

RESULT 9
BG444274
LOCUS
DEFINITION
BG444274
VERSION
BG444274.1 GI:13353926
KEYWORDS
SOURCE
ORGANISM
Gossypium arboreum
Gossypium arboreum
Eukaryota; Viridiplantae;
Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids 1; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE
1 (bases 1 to 987)
AUTHORS
Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
,D., Wood,T.C., Leslie,A. and Wilkins,T.A.
TITLE
An integrated analysis of the genetics, development, and evolution
of the cotton fiber
JOURNAL
Unpublished
COMMENT
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: TAATACGACTCACTATAGGG
High quality sequence start: 40
High quality sequence stop: 800.
Location/Qualifiers
source
1..987
/organism="Gossypium arboreum"
/mol_type="mRNA"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ea0023N11f"
/tissue_type="Fibers isolated from bolls harvested 7-10
dpa"
/lab_host="E. coli"
/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT
258 a 203 c 256 g 265 t
ORIGIN
Query Match 23.3%; Score 468.2; DB 10; Length 987;
Best Local Similarity 82.2%; Pred. No. 2.9e-52;
Matches 550; Conservative 0; Mismatches 118; Indels 1; Gaps 1;
QY 939 GTGTCGGGCGCTGGCCCTGGAGTTTCATCTTTTGTCTCTCTCTGTCGACGAGTAAATA 998

```


Db 90 GGGTGGGGCCCTTATCTTGGAGCTCTTCTCTGTTATCTTCTGGTAGTCGTGATAAGAGCA 149

QY 999 TTTTATCAAGGATATACGACACAAAGATTTTGTAGTAAACTGTGAGGACACAAAT 1058

Db 150 TTCTTCAAAGAGATATACGCGGAGATGATTTTGTAGTAAACTCTCTGGTCACAGT 209

QY 1059 CAGAGGTTTGGAGCTGAAGTGGTCATATGATTAACCGTGGATTTGGCATCTGGAGGAAATG 1118

Db 210 CAGAGGTTTGGGCTGAAGTGGTCTTATGACAAACCGGAGCTAGCATCTGGCGGAAATG 269

QY 1119 ACAACAAATTTGTTTGGGAATCAACACTCAACCCAGCTGTCTCTCAAGTACTGTGAGC 1178

Db 270 ACAACAAATTTTGTGGGAATCAACATCCACCAACCTGTCTTAAATATCTGTGACC 329

QY 1179 ACAC-AGCAGCTGTTAAAGCTATTGCAATGCTCTCTCATCTTATGAGCTTCTTGCATCT 1237

Db 330 ATACTGGCTGCGTAAAGCAATGCTTGGTCTCGCATCTGCAATGCTTCTTGCATCT 389

QY 1238 GAGGAGGAATCGAGATAGATGATTCGTTTGGGAATCAACCAACCACTCACACCTT 1297

Db 390 GGGGTGGTACAGCTGATGATGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 449

QY 1298 AGCTGATGAGCACTGGAGTCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1357

Db 450 AGCTGATGAGCACTGGAGTCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 509

QY 1358 CTAGTAAAGCACTGGAGTCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1417

Db 510 CTCGTGAGCACTGGAGTCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 569

QY 1418 TCAAGCTGGGAGCTTCTTACCGGCACTTATAGGCTTCTTATCTTCTTCTTCTTCTTCTTCT 1477

Db 570 TCAAGCTGGGAGCTTCTTACCGGCACTTATAGGCTTCTTATCTTCTTCTTCTTCTTCTTCT 629

QY 1478 GATGAGCAGATTTGATGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 1537

Db 630 GATGAGCAGATTTGATGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 689

QY 1538 CTTTCCCTTAAATCAAGATTAAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 1597

Db 690 CCAATCCCTTAAATCAAGATTAAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 749

QY 1598 ACTATCAGG 1606

Db 750 CAATTCGGG 758

RESULT 10

CB828401
LOCUS
DEFINITION LjNEST86b3r Lotus japonicus nodule library 5 and 7 week-old Lotus japonicus cDNA 5', mRNA sequence.

ACCESSION CB828401

VERSION CB828401.1

KEYWORDS EST.

SOURCE Lotus japonicus

ORGANISM

REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.
TITLE 1 (bases 1 to 528)

JOURNAL Colebatch, G., Freund, S., Trevaaskis, B. and Udvardi, M.

COMMENT Lotus japonicus root nodule ESTs: tools for functional genomics

Unpublished

Contact: Udvardi MK

Molecular Plant Nutrition

Max Planck Institute of Molecular Plant Physiology

Am Muehlenberg 1, 14476 Golm, Germany

Fax: 49 331 567 8250

Email: udvardi@mpimp-golm.mpg.de

Seq primer: T7

High quality sequence stop: 528.

Location/Qualifiers

FEATURES

source

1..528
/organism="Lotus japonicus"
/mol_type="mRNA"
/cultivar="Gifu (B-129)"
/db_xref="taxon:34305"
/dev_stage="5 and 7 week-old plants"
/clone_lib="Lotus japonicus nodule library 5 and 7 week-old"
/note="Organ: Nodule; Vector: pSPORT1, Site 1: Salt; Site 2: NotI; The library was prepared using mRNA extracted from nodules of 5 and 7 week-old Lotus plants. Nodules were induced by, and contained Mesorhizobium strain R7A."

BASE COUNT 153 a 110 c 114 g 150 t 1 others

ORIGIN

Query Match 23.1%; Score 463.4; DB 14; Length 528;
Best Local Similarity 92.9%; Pred. No. 1.9e-51;
Matches 485; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 1058 TCAGAGGTTTGGAGCTGAAGTGGTCATATGATTAACCGTGGATTTGGCATCTGGAGGAAT 1117

Db 7 TCAAGGTTTGGAGCTGAAGTGGTCATATGATTAACCGTGGATTTGGCATCTGGAGGAAT 66

QY 1118 GACAACAAAATGTTTGTGGGAATCAACACTCAACCCAGCTGTCTCTCAAGTACTGTGAG 1177

Db 67 GACAAAGATTTGTTTGTGGGAATCAACACTCAACCCAGCTGTCTCTCAAGTACTGTGAG 126

QY 1178 CACACAGCAGCTGTTAAAGCTATTGCTGCTCTCTCATCTTCTGAGCTTCTTGCATCT 1237

Db 127 CATACAGCAGCTGTTAAAGCTATTGCTGCTCTCTCATCTTCTGAGCTTCTTGCATCT 186

QY 1238 GGAGGAGGAACCTGCAGATAGATGTAATCGTTTTGGGAATCAACCAAACTCACACCTT 1297

Db 187 GGCGGTGGAAACAGCAGATCGATGTAATCGTTTTGGGAATCAACCAAACTCGCACTTA 246

QY 1298 AGCTGATGAGCAGCTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 1357

Db 247 AGCTGATGAGCAGCTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 306

QY 1358 CTAGTAAAGCAGCAGATGGGTACTCCAGAACCAAGATTTGTTGGAGATACCCCACTATG 1417

Db 307 CTAGTAAAGCAGCAGATGGGTACTCCAGAACCAAGATTTGTTGGAGATACCCCACTATG 366

QY 1418 TCAAGCTGGGAGCTTCTTACCGGCACTTATAGGCTTCTTATCTTCTTCTTCTTCTTCTTCT 1477

Db 367 TCAAGTGGGCACTCTTACAGGCCATACGTATAGAGTCTTATCTTCTTCTTCTTCTTCTTCT 426

QY 1478 GATGGACAGACTATTGTAAGTGGAGCTGGAGATGAACGCTTAGGTCTTGGAGATGTTTCT 1537

Db 427 GATGGGAGACTATTGTAAGTGGAGCTGGAGATGAACCACTTAGGTCTTGGAGATGTTTCT 486

QY 1538 CTTTCCCTTAAATCAAGATTAAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 1579

Db 487 CTTTCCCAAAATCACAGATTAAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 528

RESULT 11

CB837263

LOCUS

DEFINITION MCT003C07.16457 Ice plant lambda Uni-Zap XR expression library 5

days 0.5 M NaCl treatment, Crassulacean acid metabolism, phase IV

(5:30 PM). Mesembryanthemum crystallinum cDNA clone MCT003C07 5,

mRNA sequence.

CB837263

VERSION CB837263.1

KEYWORDS EST.

SOURCE Mesembryanthemum crystallinum (common iceplant)

ORGANISM Mesembryanthemum crystallinum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Caryophyllales; Aizoaceae; Mesembryanthemum.

REFERENCE 1 (bases 1 to 680)

AUTHORS Cushman, J.C.
 TITLE An expressed sequence tag database for the common ice plant,
 Mesembryanthemum crystallinum
 JOURNAL Unpublished
 COMMENT Contact: Cushman JC
 Department of Biochemistry
 University of Nevada
 MS200, Reno, NV 89557-0014, USA
 Tel: 775-784-1918
 Fax: 775-784-1650
 Email: jcushman@unr.edu
 PCR Primers
 FORWARD: T3 20mer
 BACKWARD: T7 21mer
 Plate: 003 row: C column: 07
 Seq primer: T3 20mer
 High quality sequence stop: 680.
 Location/Qualifiers
 1. .680
 /organism="Mesembryanthemum crystallinum"
 /mol_type="mRNA"
 /db_xref="taxon:3544"
 /clone="MCT003C07"
 /tissue type="leaf"
 /dev stage="five-week-old"
 /clone lib="ice plant Lambda Uni-Zap XR expression library
 , 5 days 0.5 M NaCl treatment, Crassulacean acid
 metabolism, phase IV (5:30 PM)."
 /notes="vector: Lambda Uni-Zap XR, Bluescript SK-; Site 1:
 EcoRI; Site 2: XhoI; Library construction was performed
 according to Stratagene's recommended protocol for the
 Lambda UniZapXR vector and cDNA synthesis kit."
 BASE COUNT 181 a 145 c 163 g 191 t
 ORIGIN
 Query Match 22.8%; Score 456.8; DB 14; Length 680;
 Best Local Similarity 81.3%; Pred. No. 1.1e-50;
 Matches 530; Conservative 0; Mismatches 122; Indels 0; Gaps 0;
 QY 925 GGGCCATCGGTTACGTGTCGGGCGCTTGGCTCGAGTTCATCTCTTTGCTCTCTGGTG 984
 DB 22 GAGCATCGCTTGGGTGGGTCTTGGCATGGAGTCTCTCTGTTGTCATCTGGTAG 81
 QY 985 ACGGGATAGAAATATTATCAACAGANATATACGACACAGAGATTTGTAGTAACT 1044
 DB 82 CCGAGATAAGATATATTCCAAAGAGATATACGTTGTGAGGATGATTATGTCAGTAAAT 141
 QY 1045 GTCAGGACAAATCAGAGGTTTGTGACTGAGTGTGCATATGATAACCGTGAGTTGGC 1104
 DB 142 AACAGGACAAATCAGAGGTTGTTGGCTGGAAGTGTCTTATGATAACCGTGAGTAGC 201
 QY 1105 ATCTGGAGGAATGACAAATAATTTGTTGTTTGGAAATCAACACTCAACCCAGCCTCTCT 1164
 DB 202 ATCAGGAGGAATGACAAATAGGCTTTTGTGTGGAACCAATCTTACTCAACCACTGTT 261
 QY 1165 CAAGTACTGTGAGCACACAGACGCTGTTAAGCTATTGCACTGTCTCTCTCATCTTCATGS 1224
 DB 262 GAAGTACTGGGAACATACAGACGCTGTTAAGCAATGCTGTGTGCGCCCATCTTCATGG 321
 QY 1225 ACTTCTTGGCATCTGGAGGAGAACTGCAGATAGATGATTTCGTTTTTGGAAATCAACACCAC 1284
 DB 322 ACTCTCGCATCTGGTGGTGTACTGCTGATCGATGATCCGTTTCTGGAATCAACACCAC 381
 QY 1285 AAATCAACCTTAGCTGTATGGACACTGGAAGTCAGTTTGCATCTTGTCTGTTGCCAA 1344
 DB 382 TAATCTCAGTTGGGATGCAATGGACACTGCGAGCTGAGTTTGAATCTTGTGGTCTAA 441
 QY 1345 AAATGTCAACGAATAGTAAACACACATGGTACTCCAGAACACAGATATTGTTTGGAG 1404
 DB 442 AATGTGATGAACTTGTGCACCAACCATGATCTCTCAAAACCAATCATCTGTTGGAG 501
 QY 1405 ATACCCCACTATGTCAAAGCTGGGACTCTTACCGGCCATCTTATAGGCTTCTCTATCT 1464

Db 502 ATAGCCTAGCATGTCAAAGTTGGCAACTCTTACAGGACACACTTACAGAGTTCTTTATCT 561
 QY 1465 TGCCATCTTCCAGATGGACAGACTATTGTAATCGAGCTGGAGATGAACCGTTAGTT 1524
 Db 562 AGCTATATCACCGGATGGCCAGACCACTTTGTAATCGGACGAGGAGATGAACACTGAGATT 621
 QY 1525 CTGGAATGTTTTCCTTCCCTTAAATCACAAGATATCTGAAAGTGAATCGGA 1576
 Db 622 CTGGAATGTTTTCCTTCCCTTAAATCACAAGATATCTGAAAGTGAATCGGA 673
 RESULT 12
 CB675428 856 bp mRNA linear EST 09-APR-2003
 LOCUS OSJNE11C18 f OSJNEe Oryza sativa (japonica cultivar-group) cDNA
 DEFINITION Clone OSJNE11C18 5', mRNA sequence.
 ACCESSION CB675428
 VERSION CB675428.1 GI:29679153
 KEYWORDS EST.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 REFERENCE 1 (bases 1 to 856)
 AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
 Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
 TITLE Large-scale identification of ESTs involved in the interaction
 between rice and Magnaporthe grisea
 JOURNAL Unpublished
 COMMENT Contact: Rod Wing
 Arizona Genomics Institute
 University of Arizona
 Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
 85721-0088, USA
 Tel: 520 626 3967
 Fax: 520 621 9288
 Email: http://genome.arizona.edu
 PCR Primers
 FORWARD: gta aaa cga cgg cca gtg
 BACKWARD: gga aac agc tat gac cat g
 Plate: 11 row: C column: 18
 Seq primer: gta aaa cga cgg cca gtg.
 FEATURES
 source
 1. .856
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="OSJNE11C18"
 /tissue_type="Leaf"
 /dev_stage="3 week"
 /lab_host="DH10B"
 /clone_lib="OSJNEe"
 /note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
 XhoI; 24 hrs after inoculation with Rice Blast (70-15)"
 BASE COUNT 193 a 214 c 241 g 208 t
 ORIGIN
 Query Match 20.3%; Score 407; DB 14; Length 856;
 Best Local Similarity 75.9%; Pred. No. 3.2e-44;
 Matches 503; Conservative 0; Mismatches 160; Indels 0; Gaps 0;
 QY 616 GGTAAAGCTCTTAGAAGTTCTCGATCGCTTATAAGGTTTGGATGCACCTGCTTT 675
 Db 194 GGGCTAGTGGCCGACAGAGATCCCTAGTCACTTATAGGTGCTGGATGCTCCGCAAT 253
 QY 676 GCAAGATGATTTTATCTGAATCTGGTAGATGCTTTCACAAATGTTGGCTGTGG 735
 Db 254 GCAGATGACTTCTACCTGAACCTTGTGATGTTGCTTCGCAATAATATCTTCAGTTGG 313
 QY 736 TTGGGTAACTGTGCTATCTCTGGAATGCTTGTAGCAGCAAGGTAACTAATATATGGA 795

Db 314 ATTGGGGAATGTGTCTACTTATGGAATGATGAGCAACAGGTCCACGATATGTGA 373
Qy 796 TTTGGGGGTTGATGATTGTGTTTCTGTTGGGCTCAACGTGGTACTCATCTTGC 855
Db 374 TTTGGGGGTTGATGACAATGTCTGTTGAGTGGGTTGGGACAGCGTGGCACTCACCTTGC 433
Qy 856 TGTGGGAACATAACATGTAAGTTCAGATTGGGATGCGAGCAAGATGCAAGAAGATAAG 915
Db 434 TGTAGGGAACAACCAAGGCAAGTTTCATGTGGATGCGACTCGATGTAGAGATAAG 493
Qy 916 ATCAATGGAGGGCCATCGGTTACGTGTGGGGGCTTGGGCTGAGTTTCATCTTTTGTG 975
Db 494 AACCATGGAAAGCCATCGGATGCGAGTAGTGTCTTTCATGGAATTCATCTTGTCTTTC 553
Qy 976 TTTCTGGTGGAGGAGATAGATATTTATCAACGAGATATACGACATATGCAAGAACTTTGT 1035
Db 554 GTCAAGCAGTGTGCAAGAGATCTCTTCAACATGATATCCGTGCCAGATGATATAT 613
Qy 1036 TAGTAAACTGTGAGGACACAAATCAGAGGTTTGTGGACTGAGTGGTTCATATGATAACCG 1095
Db 614 TACTAGACTTGTGGGATATAATCGGAGGACTGTGGGCTCAAGTGTCTTATGATAACCG 673
Qy 1096 TGAGTTGGCATCTGGAGGAATGACAAATATGTTTGTGGTGAATCAACACTCAACCCA 1155
Db 674 TCAGCTTGCATCTGGTGTATGACAAAGATTTATGTATGGAATCAACACTCGGCGCA 733
Qy 1156 GCCTGTCTCAAGTACTGTGAGCACACAGAGCTGTTTAAAGCTATTGCTGCTCTCTCA 1215
Db 734 CCGGTACTGAGTATGACTGAGCATAAGGAGCTGTCAAGTATTGCGTGGTCACTCTCA 793
Qy 1216 TCTTCATGGAATCTTGCATCTGAGGAGGAGTGTGAGATGATGATTTGTTTGGAA 1275
Db 794 TCTTCATGAGGCTGTGCTGATCTGTTGGAAGAACTGCAGATAGATGATGATTTTGGAA 853
Qy 1276 TAC 1278
Db 854 TAC 856

RESULT 13
AW030735
LOCUS EST273990 tomato callus, TAMU Lycopersicon esculentum cDNA clone
DEFINITION cLSEC25N23, mRNA sequence.

ACCESSION AW030735
VERSION AW030735.1 GI:5889491
KEYWORDS EST.

SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum

REFERENCE AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 710)
Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.
, Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Ahn,S., Ronning
, C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
Generation of ESTs from tomato callus tissue

Unpublished
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime sequence.

FEATURES
source
1..710
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLSEC25N23"
/tissue_type="callus"
/dev_stage="25-40 days old"

/lab host="XL1-Blue MRP"
/clone lib="tomato callus, TAMU"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
xhoI; supplier: Giovannoni laboratory; cLSEC - Cotyledons
of seedlings 7-10 days post-germination were excised, cut
at both ends and placed on MS medium with no selection.
Mixed callus was harvested at 25 and 40 days and included
undifferentiated masses. Tomato Callus EST Library"

BASE COUNT 209 a 142 c 149 g 209 t 1 others
ORIGIN

Query Match 19.5%; Score 392; DB 9; Length 710;
Best Local Similarity 79.5%; Pred. No. 3.4e-42;
Matches 464; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

Qy 1026 AAGATTTTGTAGTAAACTGTGAGACACAAATCAGAGGTTTGTGCACTGAAGTGGTCAT 1085
Db 1 AAGATTTTGTAGTAAACTGTGAGACACAAATCAGAGGTTTGTGCACTGAAGTGGTCAT 60

Qy 1086 ATGATAACCGTGAATAGCTTCAAGTGAATGACAAATTTGTTTGGTAATCAAC 1145
Db 61 ATGATAACCGTGAATAGCTTCAAGTGAATGACAAATTTGTTTGGTAATCAAC 120

Qy 1146 ACTCAACCCAGCTGCTCTCAAGTACTGTGAGCACACAGAGCTGTTAAAGCTATTGCA 1205
Db 121 ATTCAACACACCTTGTCTGAATACTGTGAGCATCTGCTGCGGTTAAGSCCATTCAT 180

Qy 1206 GGTCTCTCTCATCTTCATGCACTTCTGCACTGTGAGGAGGAACCTGCAGATAGATGATTC 1265
Db 181 GGTCTCTCTCATCTTCATGCACTTCTGCACTGTGAGGAGGAACCTGCAGATAGATGATTC 240

Qy 1266 GTTTTGTGAATACACCAACCAACCTTACCTGTATGACACTCGAAGTCAAGGTTT 1325
Db 241 GATTTGTGAATACACCAACCAACCTTACCTGTATGACACTCGAAGTCAAGGTTT 300

Qy 1326 GCAATCTTCTGTGGTCCAAAATGTCAACAGTATGTAAGCACACATGGGTATCCCGAGA 1385
Db 301 GCAATCTTCTGTGGTCCAAAATGTCAACAGTATGTAAGCACACATGGGTATCCCGAGA 360

Qy 1386 ACCAGTATTGTTTGGAGATACCCCACTATGTCAAGCTGGCACTCTTACCGGCCATA 1445
Db 361 ATCAGATATATGTTTGGAGATATCCGACAAATGCTTAAGATAGTACTCTGACAGGTCTATA 420

Qy 1446 CTTATAGGGTCTCTATCTTTCATCTTCCATCTCTCCAGATGGAACACTTATCAGGTGA 1505
Db 421 CATATAGAGTCTTATATCTTGTCTATCTCCAGATGGAACACTTATCAGGTGA 480

Qy 1506 GAGATGAACCGTCTAGGTTCTGGAATGTTTCCCTCCCTAAATCAAGATATCTGAAA 1565
Db 481 GAGATGAACCGTCTAGGTTCTGGAATGTTTCCCTCCCTAAATCAAGATATCTGAAA 540

Qy 1566 GTGAATATCGAGCATATCTCTTGTGAAGAACTTACTATCAGGTGA 1609
Db 541 GTGAATATCGAGCATATCTCTTGTGAAGAACTTACTATCAGGTGA 584

RESULT 14
AJ498616

LOCUS AJ498616 460 bp mRNA linear EST 29-AUG-2002
DEFINITION AJ498616 MTPOSE Medicago truncatula cDNA clone mt--acc955206907,
mRNA sequence.

ACCESSION AJ498616
VERSION AJ498616.1 GI:22089059
KEYWORDS EST.

SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula

REFERENCE 1 (bases 1 to 460)
Firnhaber,C., Bartelsemeier,V., Meyer,F., Bartels,D., Bekel,T.,
Linke,B., Puehler,A. and Kuester,H.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids 1; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.

Determination of transcript sequences from developing pods including seeds of Medicago truncatula genotype Al7		
Unpublished		
Contact:	Kuester H	
	Lehrstuhl fuer Genetik	
	Universitaet Bielefeld	
	Postfach 100131, D-33501 Bielefeld, Germany.	
Location/Qualifiers		
1.	.460	
/organism="Medicago truncatula"		
/mol_type="mRNA"		
/db_xref="taxon:3880"		
/clone="mt--acc955206g07"		
/tissue_type="pods including seeds"		
/dev_stage="different stages of development"		
/clone_lib="MTPOSE"		
/note="Vector: pGEN-T; Site 1: PetI; Site 2: SphI; genotype Al7; cDNA was prepared from polyA+ enriched RNA from developing pods including seeds harvested at different stages of development. The cDNA was directionally ligated by MediGenomix into the pGEM-T vector from Promega using GCATGGCGGAGCGGCCACATG and CTGACGCCATATGCCCCGGG adapters. Plasmids containing cDNA inserts were propagated in E. coli DH10B cells."		
BASE COUNT	131 a 108 c 83 g 138 t	
ORIGIN		
Query Match	18.8%; Score 376.4; DB 9; Length 460;	
Best Local Similarity	91.8%; Pred. No. 4.9e-40;	
Matches 423; Conservative	0; Mismatches 31; Indels 8; Gaps 2;	
QY	22 AAAGAAACTCTCTCTCTATTTCTTTCTCTCTGCAAAATTTTCGAGTAGTGTATTTT 81	
Db	1 AAAGAGAAACAAGAAATCTCTTTCTCTCTGCAATTTATCGAGTAGTGTATTTT 60	
QY	82 T-----ATAAAAAATTAATTAATTTTTTTTATATAAAGCGGTGCAAAAAATCTTT 135	
Db	61 TTTTAATAAAAAATTAATTAATTTTTTTTTTTTATATAAAAACTGAAGAGAAATTTCTTT 120	
QY	136 TACAGCGTCTCTTTTTCCCGGAAAAAAATTAACACAGCTCCGCCATGACCGAACCGG 195	
Db	121 TACAGGTTATTTTTTCCCGG--AAAAATTAACACAGCTCCGCCATGACCGAACCGG 178	
QY	196 TAATCGAAATCCACCGACTTCCACCGTCAGAGACAATTTCCACCGCTGAGCCATC 255	
Db	179 TAATCGAAATCCACCGACTTCCACCGTCGGAGATAATTTCTCCACCGCGGAGCCATC 238	
QY	256 ACCGGAGTCTCGGTCTATCTAGCCGTATGATCAACAGCAACCATTTACACCTTCACCTTC 315	
Db	239 ACCGGAGTCTCGGCATGTAGCCGTATGATCAACAGCAACCATTTACACCTTCACCTTC 298	
QY	316 TCGAACCAATCTACTCCGATAGGTTTCATTCGGAGTAGATCTGCTTCGAAATTCGCTTTGTT 375	
Db	299 TCGAACCAATCTACTCCGATAGGTTTCATTCGGAGTAGATCTGCTTCGAAATTCGCTTTGTT 358	
QY	376 TGATATCAATCTCCGACAGAGACCGGATGATAGTTCCAGCGCTTATAGCACTCTCTCT 435	
Db	359 TGATATCAATCTCCGACGGAAGACCGGATGATAGTTCCAGCGCTTATAGCACTCTCTCT 418	
QY	436 GAGAACCGGCTGTTTTCGACCGGATGTTCCCGGTCGGTTAC 477	
Db	419 TAGAACCGGCTGTTTTCGACCGGATGTTTCGAGGTCGGTTAC 460	
RESULT 15		
CA764180		
LOCUS	CA764180	
DEFINITION	AP53-Pdf_02.I21.T7.085.abi IRR1 Drought Stress Panicle Library	
	Oryza sativa (indica cultivar-group) cDNA clone C0000597.5', similar to Hypothetical 62.8 kDa trp-Asp repeats containing protein in PMCI-TFG2 intergenic region, mRNA sequence.	
ACCESSION	CA764180	
VERSION	CA764180.2 GI:27545893	
	909 bp mRNA linear EST 08-JAN-2003	

Db 362 ACATTGACGGCCATACATATAGGTATTATTTAGCCATCTCCCCAGATGGACAGACT 421
Qy 1490 ATTGTAACCTGGAGCTGGAGATGAACCGCTTAGGTTCTGGAAATGTTTCCCTTCCCTTAA 1549
Db 422 ATAGTAACTGGCGCTGGTGTGATGAACCGCTTCGGTTTGGAAACGTGTTCCATCTCCCAAG 481
Qy 1550 TCACAGAATACTGAAAGTGAAA 1571
Db 482 TCCAGAGTTCTGCACGCCTAA 503

Search completed: January 23, 2004, 21:05:55
Job time : 4362 secs